

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ХАРКІВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
імені В. Н. КАРАЗІНА

БІЗНЕС- СТАТИСТИКА

НАВЧАЛЬНО-МЕТОДИЧНИЙ ПОСІБНИК

Для студентів економічного факультету
спеціальності 051 «Економіка»

Харків – 2019

УДК 311:334.72](075.8)

ББК 65.051.520.21 я 73

К96

Рецензенти: **Третяк В. П.** – доктор економічних наук, доцент,
професор кафедри економіки та менеджменту
Харківського національного університету
імені В. Н. Каразіна;

Рахман М. С. – кандидат економічних наук,
доцент кафедри маркетингу та менеджменту
зовнішньоекономічної діяльності Харківського
національного університету імені В. Н. Каразіна

*Затверджено до друку
вченою радою економічного факультету
Харківського національного університету імені В. Н. Каразіна,
(Протокол № 2 від 21 лютого 2019 р.)*

Кущенко О. І.

К96 Бізнес-статистика: навчально-методичний посібник. Харків :
ХНУ імені В. Н. Каразіна, 2019. 116 с.

Навчально-методичний посібник сприятиме формуванню у майбутніх фахівців поглиблених теоретичних, практичних та наукових знань зі статистичної методології вивчення та аналітичної оцінки різноманітних бізнес-процесів економіки України та світу.

Для студентів вищих навчальних закладів.

УДК 311:334.72](075.8)

ББК 65.051.520.21 я 73

© Кущенко О. І., 2019

© Харківський національний університет
імені В. Н. Каразіна, 2019

ЗМІСТ

ВСТУП	4
Розділ 1. СТАТИСТИЧНА МЕТОДОЛОГІЯ ЕТАПІВ АНАЛІЗУ БІЗНЕС-ПРОЦЕСІВ	7
1.1. Структури даних та випадкові величини і закони розподілу як основа бізнес-аналітики	7
1.2. Статистичні гіпотези та дисперсійний аналіз у дослідженні бізнес-процесів	24
1.3. Практичний тренінг розділу 1	39
Розділ 2. СТАТИСТИЧНА ОЦІНКА ЕФЕКТИВНОСТІ СТАНУ РОЗВИТКУ БІЗНЕСУ В КРАЇНІ.....	45
2.1. Непараметричні методи та аналіз закономірностей для якісних даних у бізнес-аналізі	45
2.2. Статистичне оцінювання контролю якості та ефективності бізнес-процесів	62
2.3. Практичний тренінг розділу 2.....	78
РЕЗЮМЕ.....	84
ДОДАТКИ	87
1. Соціально-історический аспект розвитку бізнес-освіти на кафедрі статистики ХНУ імені В. Н. Каразіна	87
2. Sources of development of business education in Ukraine	91
3. Науково-дослідне індивідуальне завдання «Статистичний аналіз бізнес-ситуації».....	99
ЛІТЕРАТУРА	111

ВСТУП

«Статистика заповнює світ...
Підприємець, економіст, купець,
законодавець – усі користуються
статистикою».

Детев

Будь-яка діяльність у сфері економіки та бізнесу передбачає роботу з інформацією, на основі якої проводиться аналіз проблемної ситуації і приймаються управлінські рішення. Ця інформація містить різні відомості з даної проблеми, в тому числі у вигляді безлічі набору даних. Роль статистики важко переоцінити, оскільки вона є наукою і мистецтвом збору та аналізу даних, де під даними розуміється будь-який вид зареєстрованої інформації. Сукупність даних може бути великою за обсягом і мати складну структуру. У цьому випадку методи прикладного статистичного аналізу, або бізнес-статистики, – найбільш ефективний інструмент для виявлення та вивчення основних характеристик цієї сукупності. Для проведення комплексного та об'єктивного аналізу бізнес-ситуації даних може не вистачити або взагалі не бути. У таких випадках методи бізнес-статистики допоможуть найбільш ефективному напрямку збору необхідної інформації.

Мета викладання навчальної дисципліни є надання студентам поглиблених знань з теорії та практики статистичного аналізу бізнес-процесів та формування навичок щодо застосування у бізнес-аналізі сучасних програмних продуктів MS-Excel, Statistika, SPSS та інших.

Основні завдання вивчення дисципліни – вивчення теоретичних та методологічних засад статистичного аналізу бізнес-процесів, набуття вмінь застосування статистичних методів для кількісного оцінювання стану розвитку бізнесу в країні.

Згідно з вимогами освітньо-професійної програми студенти повинні досягти наступних **результатів навчання**:

знати: роль статистики у бізнесі; основні етапи статистичного аналізу бізнес-процесів; класифікації та структури даних; способи перетворення даних; основні закони розподілу випадкових величин; сутність

статистичних гіпотез та необхідність їх перевірки; критерії перевірки статистичних гіпотез для різних типів даних; основи дисперсійного аналізу; статистичні методи аналізу структури сукупності та її змін; методи статистичного оцінювання ефективності бізнес-процесів;

вміти: класифікувати дані, що характеризують конкретний бізнес-процес; перетворювати первинні дані для забезпечення їх зіставності та відповідності вимогам статистичного аналізу; формулювати нульову та альтернативну гіпотези відповідно до мети статистичного аналізу; обґрунтовувати вибір критеріїв для перевірки статистичних гіпотез; застосовувати методи одновимірного та багатовимірного дисперсійного аналізу; здійснювати порівняльний аналіз структур; здійснювати статистичне оцінювання змін у структурі сукупності; використовувати статистичні методи оцінювання ефективності бізнес-процесів; застосовувати пакети прикладних програм для розв'язання задач кількісного оцінювання бізнес-процесів.

Згідно стандарту вищої освіти за спеціальністю 051 «Економіка» (наказ Міністерства освіти і науки України № 1244 від 13.11.2018 р.) студенти повинні досягти **інтегральної компетентності**, це: «Здатність роз'язувати складні спеціалізовані задачі та практичні проблеми в економічній сфері, які характеризуються комплексністю та невизначеністю умов, що передбачає застосування теорій та методів економічної науки» [23].

Загальні компетенції: Здатність до абстрактного мислення, аналізу та синтезу. Здатність застосовувати знання у практичних ситуаціях. Здатність спілкуватися іноземною мовою [36; 37]. Навички використання інформаційних і комунікаційних технологій. Здатність до пошуку, оброблення та аналізу інформації з різних джерел. Здатність до адаптації та дій в новій ситуації. Здатність приймати обґрунтовані рішення. Здатність діяти соціально відповідально та свідомо.

Спеціальні (фахові, предметні) компетентності: Здатність виявляти знання та розуміння проблем предметної області, основ функціонування сучасної економіки на мікро-, мезо-, макро- та міжнародному рівнях. Розуміння особливостей провідних наукових шкіл та напрямів економічної науки [33]. Розуміння особливостей сучасної

світової та національної економіки, їх інституційної структури, обґрунтування напрямів соціальної, економічної та зовнішньоекономічної політики держави. Здатність застосовувати економіко-математичні методи та моделі для вирішення економічних задач [21]. Здатність застосовувати комп'ютерні технології та програмне забезпечення з обробки даних для вирішення економічних завдань, аналізу інформації та підготовки аналітичних звітів. Здатність самостійно виявляти проблеми економічного характеру при аналізі конкретних ситуацій, пропонувати способи їх вирішення. Здатність проводити економічний аналіз функціонування та розвитку суб'єктів господарювання, оцінку їх конкурентоспроможності. Здатність поглиблено аналізувати проблеми і явища в одній або декількох професійних сферах з урахуванням економічних ризиків та можливих соціально-економічних наслідків [21].

Комунікація: донесення до фахівців і нефахівців інформації, ідей, проблем, рішень та власного досвіду в галузі професійної діяльності, а також здатність ефективно формувати комунікаційну стратегію.

Автономія та відповідальність: управління комплексними діями або проектами, відповідальність за прийняття рішень у непередбачуваних умовах.

Відповідальність за професійний розвиток окремих осіб або груп осіб, здатних до подальшого навчання з високим рівнем автономності (Див. Додаток 3).

Навчально-методичний посібник складається з двох розділів, кожний з яких містить основні блоки питань розділу, враховуючи самоконтроль, а також практичний тренінг, який представлено у вигляді тестів для самостійної роботи над темами курсу. В структуру посібника входить також резюме, яке підводить студента до розуміння аналітичних можливостей статистики, та три додатка, котрі дають змогу зрозуміти роль кафедри статистики у бізнес-освіті на світовому рівні. Для зручності користування додатки викладено трьома мовами.

Розділ 1

СТАТИСТИЧНА МЕТОДОЛОГІЯ

ЕТАПІВ АНАЛІЗУ БІЗНЕС-ПРОЦЕСІВ

1.1. Структури даних та випадкові величини і закони розподілу як основа бізнес-аналітики

Структури даних.

Класифікація, розподіл та перетворення даних.

*Основні переваги статистичних обстежень
бізнес-активності підприємств.*

Випадкові величини і закони розподілу як основа бізнес-статистики.

Структури даних

Статистичні дані можуть бути представлені у різних формах. Набори даних мають в собі одне чи декілька значень для кожного окремого об'єкту. В якості таких об'єктів можуть виступати люди, міста, комп'ютери, книги і все, що має інтерес для вивчення. Такі об'єкти називають *елементарними одиницями*. Для кожного об'єкта реєструють одні й ті самі ознаки. Наприклад, реєструють ріст і масу людей; чисельність населення, рівень народжуваності і смертності для окремих міст; об'єм пам'яті, швидкість роботи комп'ютеру; рік видання, кількість сторінок, видавництво книг і т. п. Ознаку, що реєструється для кожного з об'єктів, називають *змінною*.

Класифікація, розподіл та перетворення даних

Набори даних класифікують за наступними ознаками:

- за кількістю змінних (одновимірні, двовимірні, багатовимірні);
- за типом даних (кількісні, якісні);
- за тим, чи потрібна впорядкованість даних у часі або ні.

Одновимірні набори даних містять тільки одну ознаку для кожного об'єкта. Ці дані дозволяють визначити типове значення ознаки, наскільки значення відрізняються один від одного, а також дозволяють зрозуміти, чи потребують окремі дані особливої уваги. Прикладом одновимірних

даних є інформація про середню зарплату у регіоні по галузях. Вона дозволяє: назвати галузь з найвищим рівнем зарплати; зрозуміти, наскільки відрізняються рівні середніх зарплат у різних галузях один від одного; звернути увагу на галузі, де рівень зарплати найнижчий.

Набори *двовимірних* даних містять інформацію про дві ознаки для кожного з об'єктів. Крім того, що вони дають можливість отримати два набори *одновимірних* даних, двовимірні дані дозволяють встановити, чи існує зв'язок між двома змінними, наскільки сильно пов'язані змінні, чи можна передбачити значення однієї змінної за значенням іншої і якщо так, то з якою надійністю. Наприклад, дані опитування студентів про те, чи задоволені вони рівнем теоретичної і практичної підготовки, отриманої у вузі (значення обох змінних записуються у вигляді так / ні, або 1 / 0). Вони дозволяють встановити, чи є зв'язок між рівнями теоретичної та практичної підготовки у вузі.

Багатовимірні дані містять інформацію про три або більше ознаки для кожного об'єкта. На додаток до тієї інформації, яку можна отримати з *одновимірних* і *двовимірних* наборів, багатовимірні дані можна використовувати для отримання інформації про те, чи існує проста залежність між цими ознаками, наскільки вони взаємопов'язані (мова йде не тільки про попарний взаємозв'язок ознак, але і про залежність у сукупності), чи можна передбачити значення однієї змінної на підставі значень інших. Прикладом багатовимірних даних може бути число іноземних громадян, які навчаються в ХНУ на певну дату, по країнам.

Значення змінних, які реєструються за допомогою чисел, що мають змістовний сенс, називають *кількісними* даними. З *кількісними* даними можна виконувати всі звичайні операції над числами, такі як обчислення середньої та оцінку мінливості. Залежно від того, які значення може потенційно приймати змінна, виділяють два типи *кількісних* даних: *дискретні* і *безперервні*.

Дискретна – це така змінна, яка може набувати значень тільки з деякого списку певних чисел. Прикладами дискретної змінної є число дітей у сім'ї; число викликів «швидкої допомоги», що надходять до лікарні; число клієнтів, які звернулися у фірму за певний проміжок часу, і т. п.

Безперервною будемо вважати будь-яку змінну, яка не є дискретною. Вона приймає значення з деякого проміжку. Прикладами безперервної змінної є зріст дорослої людини (наприклад від 140 до 230 см); фактична маса буханки хліба (наприклад, від 750 до 830 г); дальність польоту снаряду; врожайність культури, вирощеної у господарстві, і т. п.

Є дані, які реєструють певну якість, яким володіє об'єкт. Такі дані називають *якісними*. Навіть якщо значенням цієї якості можна приписати числа (наприклад, статі людини приписати відповідно числа 0 і 1), то обробляти ці числа як кількісні дані не можна. Прикладами якісних даних є тип школи, де навчається дитина (ліцей, гімназія, спеціалізована фізико-математична школа); посаду, яку займає співробітник на підприємстві; назви газет, які читають у певному місті, і т. п.

Якісні дані бувають двох типів: *порядкові*, для яких існує змістовний сенс порядок, і *номінальні*, для яких немає змістовного порядку.

Порядкові дані можна ранжувати і використовувати. Це ранжування при проведенні статистичного бізнес-аналізу. Прикладом порядкових даних є відповіді на питання анкети, що містять наступні варіанти відповідей: так; більше так, аніж ні; більше ні, ніж так; немає. Хоч їх і можна виразити (ці відповіді числами – наприклад, 4, 3, 2, 1), але запропонована шкала оцінок носить суб'єктивний характер. Не можна вважати, що різниця між відповідями 4 і 3 така ж, як і між відповідями 2 і 1. Також не можна вважати, що відповідь 3 в три рази краще відповіді 1.

Для *номінальних* даних немає числових значень і немає основи для ранжирування. Прикладами номінальних даних є регіони України, з яких приїхали студенти ХНУ; назви фірм, що виготовляють миючі засоби; стать працівників фірми, і т. п.

Якщо порядок запису значень даних у часі має змістовний сенс, то кажуть, що ці дані являють собою *часовий ряд*. Ці дані представляють інформацію про об'єкт у різні моменти часу. Якщо порядок запису даних у часі не суттєвий, то говорять про один часовий зріз. Ці дані представляють інформацію про об'єкти у певний момент часу. Прикладами часового ряду є дані про: злочинність у регіоні за кілька років; рівень безробіття за кілька років; динаміку успішності учнів

з математики і т. п. Прикладами одного часового зрізу є дані про кількість злочинів у районах області в певний рік; число безробітних серед різних вікових груп населення за даними на певний день; дані про вірне виконання тестів учнями класу і т. п.

Іноді дані класифікують також за тим, чи збиралися вони спеціально для запланованого аналізу або збиралися раніше для інших потреб. У першому випадку їх називають *первинними даними*, а у другому – *вторинними даними*. Отримання первинних даних пов'язане найчастіше з великими витратами коштів та часу, але зате дослідник отримує те, що йому потрібно. Вторинні дані, як правило, обходяться дешевше, їх швидше можна отримати, але при цьому не завжди в них можна знайти те, що потрібно.

Багато статистичних даних отримують у процесі вимірів. Метою вимірювань є отримання інформації про ознаки об'єктів, організмів, подій. Вимірюється не сам об'єкт, а тільки властивості або відмінні ознаки об'єкта. Наприклад, вимірюється не учень, а його зростання, маса, швидкість читання, досягнення з математики, спортивні досягнення і т. і. Вимірювання здійснюються шляхом встановлення відповідності між числами та об'єктами, які є носіями та підлягають вимірюванню властивостей. Вимірювання можуть проводитися на різних рівнях. Різним рівням вимірювань відповідають різні шкали:

- номінальна шкала;
- порядкова (рангова);
- шкала інтервалів;
- шкала відношень (шкала пропорцій).

Номінальна шкала використовується для реєстрації найнижчого рівня вимірювань, який передбачає наявність мінімальних передумов для вимірювання. При вимірюваннях на даному рівні практично не використовуються числа. Тут важливо встановити подібність або відмінність об'єктів за певною ознакою, тобто, при цьому мають справу з якісними даними. Розглянемо приклади:

1. Розподіл учнів по класах за: статевою ознакою, місцем проживання, видами спорту, якими вони займаються, числом дітей у сім'ї є прикладами величин номінальної шкали. При цьому

можливий розподіл учнів за двома або більше ознаками (двовимірні або багатовимірні дані).

2. Перелік фірм, що займаються виробництвом: вантажних та легкових автомобілів, автомобілів спеціального призначення і автобусів. Такі відмінні ознаки вироблених автомобілів є прикладами величин номінальної шкали.

За допомогою підрахунку можна встановити частоту тієї чи іншої категорії (число хлопчиків та дівчат у школі; число учнів, які проживають у кожному мікрорайоні; число учнів у кожному класі; число учнів, що займаються тим чи іншим видом спорту; кількість фірм, що займаються виробництвом автобусів і т. п.). При цьому можна визначити, яка величина найбільш часто зустрічається (клас, в якому навчається найбільша кількість учнів; вид спорту, який користується найбільшою популярністю в учнів; тип автомобіля, виробництвом якого займається найбільше число фірм). Категорії даних номінальної шкали позначаються, як правило, словесно (вербально).

Порядкова (рангова) шкала вказує лише послідовність носіїв ознаки або напрямок ступеня вираженості ознаки. Наприклад, учнів можна ранжувати за кількістю правильно виконаних тестових завдань. Нехай учні А, Б, В, Г, Д правильно виконали відповідно 21, 16, 12, 9 і 3 завдання. Графічно це можна зобразити так:

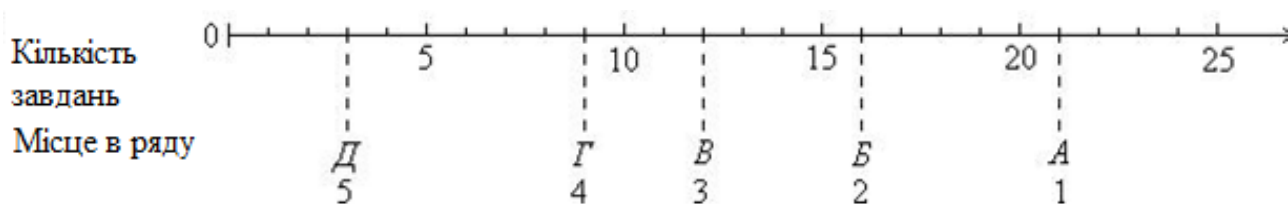


Рис. 1.1.1. Кількість вірно виконаних учнями завдань

Ця порядкова шкала має величини від 1 до 5, і учні на ній розміщені в залежності від кількості правильно виконаних завдань: А – перший, Д – п’ятий. З рис. 1.1.1 видно, що інтервали, що розділяють місця в ряду, різні за величиною. З цієї причини недоцільно складати, віднімати, множити та ділити порядкові місця.

Шкала шкільних оцінок з одного предмета є порядковою шкалою, так як інтервали між окремими балами не відображають розриву між

реальними результатами. Ми знаємо тільки, що учень, який отримав оцінку «5» з якогось предмету, знає цей предмет краще того, хто отримав «4». Але не можна стверджувати, що різниця у знаннях цих учнів така ж, як і в знаннях тих, хто отримав «4» та «3». Так як шкала оцінок є порядковою шкалою, то некоректно виставляти підсумкову оцінку як середню арифметичну поточних оцінок.

Іншим прикладом порядкової шкали є шкала посад, які займають працівники на підприємстві: директор, заступник директора, начальник відділу, завідувач лабораторією, старший інженер, інженер, молодший інженер, технік, старший лаборант, лаборант.

На шкалі *інтервалів* рівні інтервали відображають однакову міру величини вимірюваної ознаки. Наприклад: 1 см між 3-м і 4-м сантиметрами на шкалі вимірювань довжин має такий же зміст, як і 1 см між 82-м і 83-м сантиметрами. Іншими словами, на шкалі інтервалів відстані між сусідніми поділами рівні. При користуванні інтервальною шкалою цілком осмисленим є питання «На скільки?», але не завжди, користуючись цією шкалою, можна формулювати питання «У скільки разів?». Справа у тому, що на шкалі інтервалів встановлюються довільно початок відліку (нуль шкали), одиниця виміру та напрямок відліку. Прикладом інтервальної шкали є температурна шкала за Цельсієм. Різниця між температурами повітря $+30^{\circ}\text{C}$ і $+20^{\circ}\text{C}$ настільки ж велика, як і між -10°C і -20°C . Однак не можна стверджувати, що при температурі повітря $+30^{\circ}\text{C}$ у півтора рази тепліше, ніж при температурі $+20^{\circ}\text{C}$. Навіть якщо температура повітря дорівнює 0°C , не можна стверджувати, що тепла немає зовсім: адже початок відліку вибрано довільно. Шкали на більшості фізичних приладів (амперметр, вольтметр і ін.) є інтервальними.

Шкала коефіцієнта інтелекту IQ є шкалою інтервалів. Шкала інтервалів є метричною, з її допомогою можна виконувати додавання і віднімання. Вона має значні переваги в порівнянні з номінальною і порядковою шкалами.

Шкала відношень, або шкала пропорцій, дає можливість встановлювати відносини значень вимірюваної ознаки завдяки тому, що значенням шкали «0» відповідає величина, для якої ознака відсутня.

Іншими словами, початок відліку на цих шкалах вибирають мимоволі. Прикладами шкали відносин є міри довжини (м, см і т. п.) та маси (кг, г і т. п.). Предмет довжиною 100 см удвічі довше предмета довжиною 50 см.

Іноді дані потребують перетворення. Зокрема, потреба у цьому виникає, коли у ряду даних одне або кілька даних істотно перевищують інші. Якщо дані явно несиметричні, то кожне значення наведеного набору даних замінюють логарифмом цього значення з метою спростити статистичний бізнес-аналіз.

Основні переваги статистичних обстежень бізнес-активності підприємств

Функціонування будь-якого суб'єкта господарювання на ринку не може відбуватися успішно без постійного використання інформації про: власний стан, стан навколишнього середовища, партнерів, конкурентів, що й сприяло формуванню нового напрямку у практиці національної статистики – статистичного моніторингу. *Моніторинг* (від лат. monitor – що нагадує, що наглядає) – у широкому розумінні є функцією управління поряд з функціями планування, стратегічного та операційного менеджменту. Удосконалення управління зумовлює використання менеджерами результатів статистичного моніторингу ділової активності підприємств для розробки подальших можливих корпоративних стратегій розвитку: реструктуризації, диверсифікації, вертикальної інтеграції та портфельного аналізу.

Ці корпоративні стратегії розвитку досить вичерпно схарактеризовані у відповідних літературних джерелах. Зазначимо тут лише, що *портфельним* називають аналіз, для проведення якого потрібне статистичне забезпечення у вигляді статистичного моніторингу ділової активності підприємств. Статистичний моніторинг ділової активності підприємств має цілком визначену мету, алгоритм функціонування і використовує систему якісних та кількісних статистичних показників.

Традиційно портфельний аналіз використовується з метою оптимізації портфеля цінних паперів інвестора. Фінансист або статистик під терміном «*портфоліо*» розуміє оптимальний з погляду комбінації ризику та прибутковості набір інвестицій. Отже, в основу портфельного аналізу

покладено два оцінні критерії: теперішня вартість очікуваних доходів від володіння цінними паперами (проценти, дивіденди) та рівень ризикованості вкладень.

За аналогією з цінними паперами можна проводити аналіз портфеля продуктів (товарів та послуг) підприємства. Це означає, що портфельний аналіз доцільний на тих підприємствах, які виробляють багато видів продуктів. У процесі аналізу окремі групи продуктів розглядаються як відповідні стратегічні «бізнес-одиниці», причому кожна з них оцінюється щодо прибутковості та ризикованості виробництва. На підставі результатів портфельного аналізу приймаються рішення про додаткові інвестиції в окремі виробничі програми, реінвестиції чи дезінвестиції, а також вивчається стратегія управління ризиками за кожною зі стратегічних бізнес-одиниць.

Під час бізнес-аналізу використовуються чотири широковідомі у теорії та практиці маркетингу портфельні матриці:

- матриця «зростання ринку – частка ринку» (матриця Бостонської консультативної групи);
- матриця «привабливість ринку – конкурентні переваги» (матриця Мак-Кінсі);
- матриця General Electric «конкурентоспроможність стратегічних бізнес-одиниць – привабливість галузі»;
- матриця напрямленої політики «перспективи розвитку ділового сектору (погані, середні або добрі) – конкурентоспроможність суб'єктів господарювання (низька, середня або висока)».

Саме остання, четверта матриця – матриця напрямленої політики – має за систему координат перспективи розвитку ділового сектору. Тому результати статистичного моніторингу ділової активності підприємств мають бути спрямовані саме на агреговану оцінку перспектив розвитку ділового сектору.

Згідно з планом статистичних робіт Державна служба статистики України проводить регулярні щоквартальні обстеження ділової активності підприємств України, за результатами яких видається Бюлетень обстежень ділової активності підприємств країни. Обстеження

ділової активності охоплює промислові, будівельні, торговельні, сільськогосподарські та транспортні підприємства.

В Україні щомісячно опитується 1313 промислових підприємств, 244 будівельних, 726 торговельних, 590 сільськогосподарських та 251 транспортне підприємство. Щодо причин, які спонукають респондентів до участі в опитуваннях, то 37 % від загальної їхньої кількості вказало на те, що обстеження дають привід обмірковувати стан підприємства, майже кожний п'ятий із опитуваних отримує корисну інформацію в обмін на участь в обстеженнях.

Які ж основні переваги обстежень ділової активності? По-перше, – це швидкість отримання необхідної інформації. Обстеження за наявності комп'ютерних технологій обробки роблять доступними ключові дані протягом 15–20 днів, тоді як офіційні цифри отримуються значно пізніше. По-друге, інформація з обстежень доповнює дані офіційної статистики, що дає змогу краще і швидше показувати поширення економічних тенденцій. По-третє, інформація, яку отримують за допомогою обстежень є суб'єктивною, так як ці результати вимірюють розуміння ділової ситуації респондентами – менеджерами підприємств. Наприклад, за результатами обстеження легко отримати інформацію про очікування підприємців і виявити на її підставі не тільки їхнє сприйняття економічної ситуації, а й спрогнозувати їхню економічну поведінку. І остання, найбільш очевидна перевага обстежень: вони дозволяють отримувати інформацію зі сфер економічної діяльності, що не охоплені кількісною офіційною статистикою, наприклад, про наявність та використання виробничих потужностей, оцінку запасів готової продукції та ін. Обстеження проводяться за методикою Комісії Європейського союзу, рекомендованою для застосування у країнах з перехідною економікою.

Випадкові величини і закони розподілу як основа бізнес-статистики

Змінна величина називається *випадковою*, якщо у результаті досвіду вона може приймати дійсні значення з певними ймовірностями. *Закон розподілу* – функція (таблиця, графік, формула), що дозволяє визначати

ймовірність того, що випадкова величина X приймає певне значення x_i або потрапляє у певний інтервал.

Випадкова величина X називається дискретною, якщо існує така невід’ємна функція:

$$P\{X = x_i\} = p_i, \quad i = 1, 2, \dots, \quad \sum_{i=1}^{\infty} p_i = 1 \quad (1.1.1),$$

яка ставить у відповідність значенням x_i змінної X ймовірність p_i , з якої вона приймає це значення.

Випадкова величина X називається *безперервною*, якщо для будь-яких $a < b$ існує така невід’ємна функція $f(x)$:

$$P\{a \leq X \leq b\} = \int_a^b f(x) dx \quad (1.1.2).$$

Функція $f(x)$ називається *щільністю розподілу* неперервної випадкової величини. Ймовірність того, що випадкова величина X (дискретна або безперервна) приймає значення, менше x , називається *функцією розподілу* випадкової величини X і позначається $F(x)$:

$$F(x) = P\{X < x\} \quad (1.1.3).$$

Кожен закон розподілу – це деяка функція, що повністю описує випадкову величину з ймовірнісної точки зору. На практиці про розподіл ймовірностей випадкової величини X часто доводиться судити лише за результатами випробувань. Повторюючи випробування, будемо кожен раз реєструвати, чи відбулася випадкова подія A , чи ні. Відносною частотою (або просто частотою) випадкової події A називається відношення числа n_A появ цієї події до загального числа n проведених випробувань.

Проводячи серії випробувань, які дають емпіричні значення величини X , треба фіксувати числа n_x влучень результатів у кожний інтервал. При великому числі випробувань n ставлення n_x/n (частоти потрапляння в інтервали) повинні бути близькі до можливості попадання у ці інтервали. Залежність частот n_x/n від інтервалів визначає *емпіричний*

розподіл ймовірностей випадкової величини X , графічне представлення якої називається гістограмою (рис. 1.1.2).

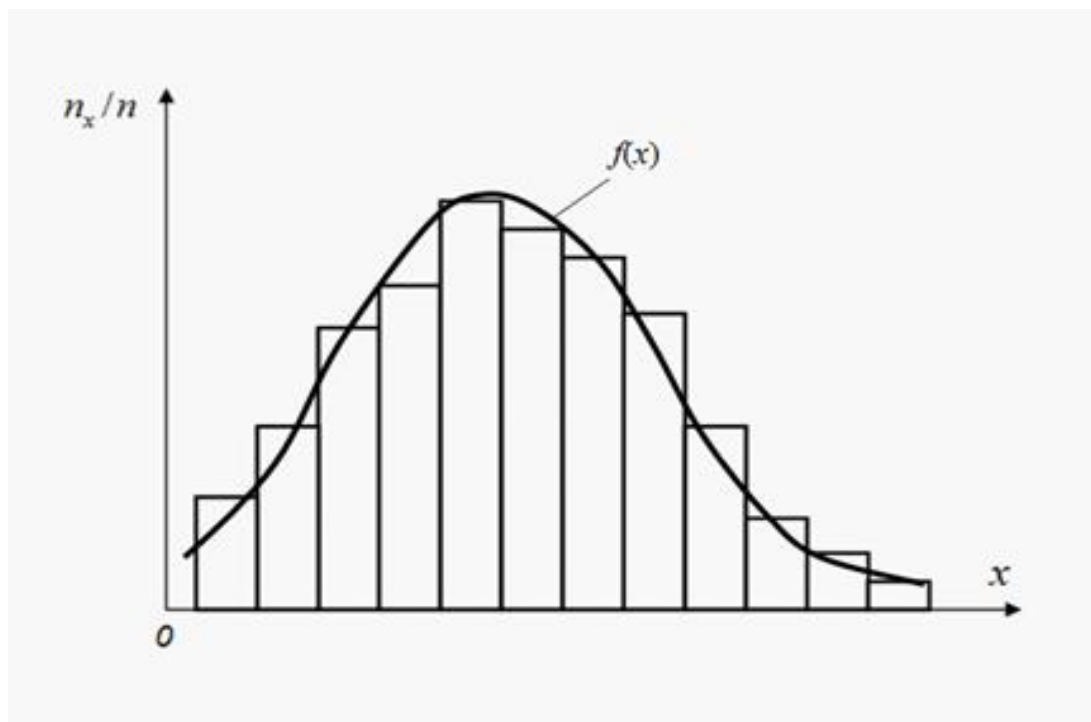


Рис. 1.1.2. Гістограма та щільність розподілу

Для побудови гістограми по осі абсцис відкладають інтервали рівної довжини, на які розбивається весь діапазон можливих значень випадкової величини X , а по осі ординат відкладають частоти n_x/n . Тоді висота кожного стовпчика гістограми дорівнює відповідній частоті. Таким чином, виходить наближене уявлення закону розподілу ймовірностей для випадкової величини X у вигляді ступінчастої функції, апроксимація (вирівнювання) якої деякої кривої $f(x)$ дасть щільність розподілу.

Розподіл дискретних випадкових величин

Біноміальний розподіл. Дискретна випадкова величина X має біноміальний розподіл, якщо її можливі значення $0, 1, 2, \dots, m, \dots, n$, а відповідні їм ймовірності рівні:

$$P_m = P\{X = m\} = C_n^m p^m q^{n-m} \quad (1.1.4).$$

Як бачимо з (1.1.4), ймовірності P_m обчислюються як члени розкладання бінома Ньютона $(p+q)^n$, звідки і назва «біноміальний розподіл». Прикладом є вибірковий контроль якості виробничих виробів, при якому відбір виробів для проби проводиться за схемою *випадкової*

повторної вибірки, тобто коли перевірені вироби повертаються у вихідну партію. Тоді кількість нестандартних виробів серед відібраних є випадкова величина з біноміальним законом розподілу ймовірностей. Біноміальний розподіл визначається двома параметрами: n і p . Випадкова величина, розподілена за біноміальним законом, має наступні основні числові характеристики:

$$m = np, \quad D = npq, \quad \sigma = \sqrt{npq} \quad (1.1.5).$$

Геометричний розподіл. Дискретна випадкова величина X має геометричний розподіл, якщо її можливі значення $0, 1, 2, \dots, m, \dots, n$, а ймовірності цих значень:

$$m = np, \quad D = npq, \quad \sigma = \sqrt{npq} \quad (1.1.6),$$

де: $0 < p < 1, q = 1 - p; m = 0, 1, 2, \dots$

Розподіл неперервних випадкових величин

Рівномірний розподіл. Величина X розподілена рівномірно в інтервалі (a, b) , якщо всі її можливі значення знаходяться у цьому інтервалі та щільність розподілу ймовірностей постійна:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{при } x \in (a, b), \\ 0 & \text{при } x \notin (a, b). \end{cases} \quad (1.1.7)$$

Для випадкової величини X , рівномірно розподіленою в інтервалі (a, b) (рис. 1.1.3), ймовірність попадання в будь-який інтервал (x_1, x_2) , що лежить всередині інтервалу (a, b) , дорівнює:

$$P\{x_1 < X < x_2\} = \frac{x_2 - x_1}{b - a} \quad (1.1.8)$$

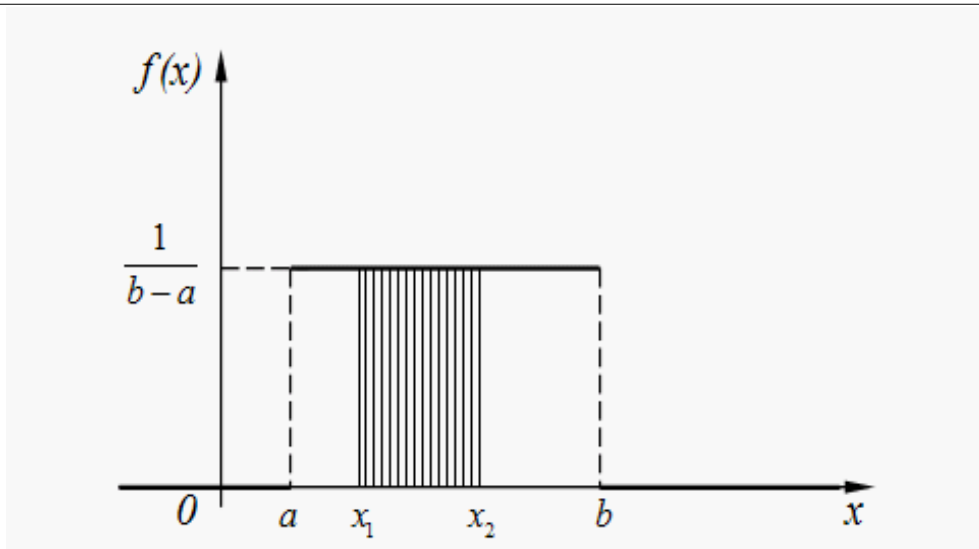


Рис. 1.1.3. Щільність рівномірного розподілу

Прикладами рівномірно розподілених величин є помилки округлення. Так, якщо усі табличні значення деякої функції округлені до одного і того ж розряду 10^{-m} , то вибираючи навмання табличне значення, ми вважаємо, що помилка округлення обраного числа є випадкова величина.

Показовий розподіл. Безперервна випадкова величина X має показовий розподіл, якщо щільність розподілу її ймовірностей виражається формулою:

$$f(t) = \begin{cases} \lambda * e^{-\lambda t} & \text{при } t > 0, \\ 0 & \text{при } t \leq 0. \end{cases} \quad (1.1.9)$$

Графік щільності розподілу ймовірностей (1.1.9) представлено на рис. 1.1.4.

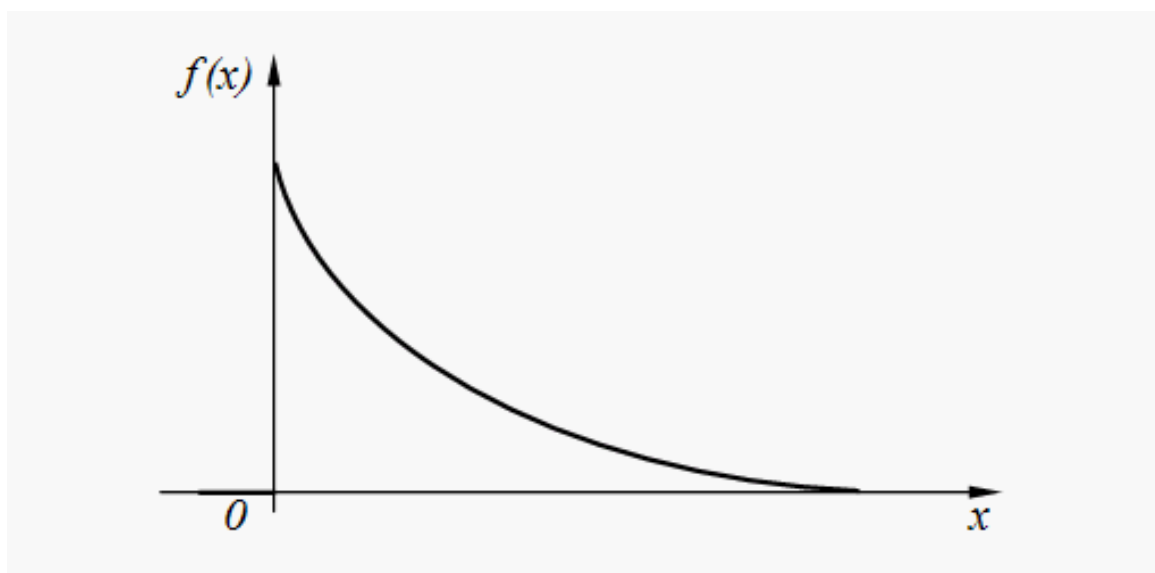


Рис. 1.1.4. Щільність показового розподілу

Нормальний (гауссовий) розподіл. Випадкова величина X має нормальний (гауссовий) розподіл, якщо щільність розподілу її ймовірностей визначається залежністю:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}, \quad (1.1.10)$$

де: $m = M(X)$, $\sigma = \sqrt{D(X)}$.

При $m = 0$, $\sigma = 1$ нормальний розподіл називається *стандартним*.

Графік щільності нормального розподілу (1.1.10) представлений на рис. 1.1.5.

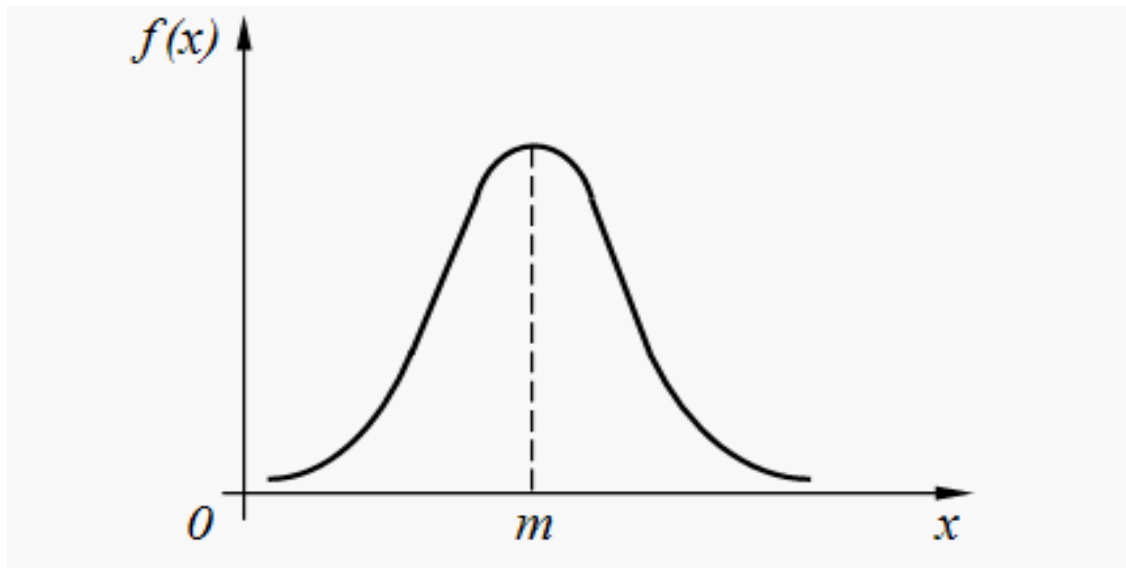


Рис. 1.1.5. Щільність нормального розподілу

Нормальний розподіл є таким, який найбільш часто зустрічається у різних випадкових явищах природи. Так, помилки виконання команд автоматизованим пристроєм, помилки виведення космічного корабля у задану точку простору, помилки параметрів комп'ютерних систем і т. п. у більшості випадків мають нормальне або близьке до нормального закону розподілу. Більш того, випадкові величини, утворені підсумовуванням великої кількості випадкових даних, розподілені практично за нормальним законом.

Системи випадкових величин

Істотний інтерес у математичній статистиці представляє розгляд системи двох та більше випадкових величин, а їх статистичний взаємозв'язок один із одним. За аналогією з рядом розподілу однієї

дискретної величини X для двох дискретних випадкових величин X і Y будується матриця розподілу – прямокутна таблиця, у якій записані всі ймовірності: $P_{ij} = P\{X = x_i, Y = y_j\}$, $i = 1, \dots, n$; $j = 1, \dots, m$. То події (або досліди), які називаються незалежними, кажуть, що ймовірність появи (результату) кожного з них не залежить від того, які події (результати) мали місце у інших випадках (подіях або дослідах). Дві випадкові величини X і Y називаються незалежними, якщо незалежні всі пов'язані з ними події: наприклад, $\{X < a\}$ і $\{Y < b\}$ або $\{X = x_i\}$ і $\{Y = y_j\}$ і т. п. Спільною функцією розподілу системи двох випадкових величин (X, Y) називається ймовірність спільного виконання нерівностей $X < x$ і $Y < y$:

$$F(x, y) = P\{X < x, Y < y\}. \quad (1.1.11)$$

Подія означає твір (спільне виконання) подій $\{X < x\}$ і $\{Y < y\}$.

Геометричною інтерпретацією спільної функції розподілу $F(x, y)$ є ймовірність попадання випадкової точки (X, Y) на площині всередину квадранта з вершиною в точці (x, y) (заштрихована область на рис. 1.1.6).

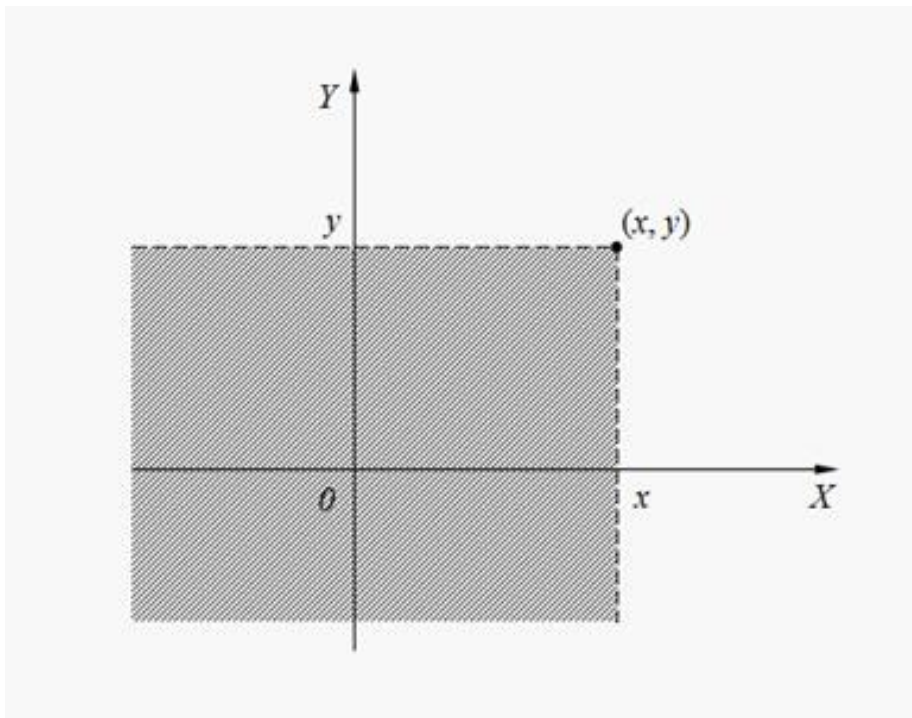


Рис. 1.1.6. Геометрична інтерпретація спільної функції розподілу $F(x, y)$

В якості числових характеристик системи двох випадкових величин X і Y зазвичай розглядаються початкові та центральні моменти різних порядків. Порядком моменту називається сума його індексів $k + s$.

Початковим моментом порядку $k + s$ системи двох випадкових величин X і Y називається математичне очікування добутку X^k на Y^s :

$$\alpha_{k,s} = M(X^k Y^s). \quad (1.1.12).$$

Центральним моментом порядку $k + s$ системи двох випадкових величин (X, Y) називається математичне сподівання добутку $(X - m_x)^k$ на $(Y - m_y)^s$:

$$\mu_{k,s} = M[(X - m_x)^k (Y - m_y)^s], \quad (1.1.13)$$

де: $m_x = M(X)$, $m_y = M(Y)$.

В інженерному застосуванні математичної статистики найчастіше використовуються моменти першого та другого порядків.

Початкові моменти першого порядку

$$\alpha_{1,0} = M(X^1 Y^0) = M(X) = m_x; \quad \alpha_{0,1} = M(X^0 Y^1) = M(Y) = m_y \quad (1.1.14)$$

є математичними очікуваннями випадкових величин X і Y .

Центральні моменти першого порядку завжди дорівнюють нулю:

$$\mu_{1,0} = M[(X - m_x)^1 (Y - m_y)^0] = M(Y - m_x) = M(X) - M(m_x) = m_x - m_x = 0$$

$$\mu_{0,1} = M[(X - m_x)^0 (Y - m_y)^1] = M(Y - m_y) = M(Y) - M(m_y) = m_y - m_y = 0 \quad (1.1.15)$$

Початкові моменти другого порядку:

$$\alpha_{2,0} = M(X^2 Y^0) = M(X^2) = \alpha_2(X),$$

$$\alpha_{0,2} = M(X^0 Y^2) = M(Y^2) = \alpha_2(Y),$$

$$\alpha_{1,1} = M(X^1 Y^1) = M(XY) \quad (1.1.16).$$

Центральні моменти другого порядку:

$$\mu_{2,0} = M[(X - m_x)^2 (Y - m_y)^0] = M[(X - m_x)^2] = D_x,$$

$$\mu_{0,2} = M[(X - m_x)^0 (Y - m_y)^2] = M[(Y - m_y)^2] = D_y,$$

$$\mu_{1,1} = M[(X - m_x)^1 (Y - m_y)^1] = M[(X - m_x)(Y - m_y)]. \quad (1.1.17)$$

де: D_x, D_y – дисперсії випадкових величин X і Y .

Центральний момент другого порядку називається коваріацією випадкових величин X і Y . Позначимо його:

$$K_{x,y} = M[(X - m_x)(Y - m_y)]. \quad (1.1.18)$$

З визначення коваріації (1.1.18) слідує:

$$K_{xy} = K_{yx} \quad (1.1.19)$$

Дисперсія випадкової величини є по суті окремим випадком коваріації:

$$D_x = M[(X - m_x)(X - m_x)] = K_{xx},$$

$$D_y = M[(Y - m_y)(Y - m_y)] = K_{yy} \quad (1.1.20)$$

Коваріація двох випадкових величин X і Y характеризує ступінь їх залежності та міру розсіювання навколо точки. Часто буває зручно виразити ковариацию у вигляді:

$$K_{xy} = M(X * Y) - M(X) * M(Y) \quad (1.1.21)$$

Розмірність ковариації дорівнює добутку розмірностей випадкових величин X і Y . Безрозмірна величина, що характеризує тільки залежність випадкових величин X і Y :

$$r_{xy} = \frac{K_{xy}}{\sigma_x \sigma_y} \quad (1.1.22)$$

називається *коефіцієнтом кореляції* випадкових величин X і Y . Цей параметр характеризує ступінь лінійної залежності випадкових величин X і Y . Для будь-яких двох випадкових величин X і Y коефіцієнт кореляції $-1 \leq r \leq 1$. Якщо $r > 0$, то лінійна залежність між X і Y зростаюча, якщо,

$r < 0$ то лінійна залежність між X і Y зворотня, при $r = 0$ лінійної залежності між X і Y немає. Відсутність лінійної кореляції не означає відсутність будь-якої іншої залежності між X і Y . Якщо має місце жорстка лінійна залежність $Y = a_x + b$, то $r = +1$ при $a > 0$ і $r = -1$ при $a < 0$.

Контрольні питання

1. За якими критеріями класифікують дані? Наведіть приклади.
2. Назовіть шкали вимірювань на різних рівнях.
3. Основні переваги статистичних обстежень бізнес-активності підприємств.
4. Поняття про «портфоліо» та матриці портфельного аналізу.
5. Що таке випадкова величина и чим вона відрізняється від числа?
6. Вкажіть відмінності між дискретною та неперервною випадковою величиною.
7. Системи випадкових величин та їх графічна інтерпретація.
8. Поняття про «коваріацію» та її характеристика.

1.2. Статистичні гіпотези

та дисперсійний аналіз у дослідженні бізнес-процесів

Поняття статистичної гіпотези та методика її перевірки.

Критерії перевірки статистичних гіпотез, типи та їх властивості.

Статистичний висновок як основа узагальнення даних.

Дисперсійний аналіз у дослідженні бізнес-процесів

Поняття статистичної гіпотези та методика її перевірки

Статистична гіпотеза – певне припущення про властивості розподілу ймовірностей, що лежить в основі спостережуваних випадкових явищ. Результати спостережень представляються зазвичай у вигляді реалізації деякої сукупності випадкових величин. При цьому спільний розподіл цих випадкових величин відомий у повному обсязі і передбачає належність його до деякого певного класу розподілів. У такій ситуації ставиться завдання статистичних гіпотез перевірки.

Перевірка статистичної гіпотези – це процес прийняття рішення про те, чи суперечить розглянута статистична гіпотеза, яка спостерігається, вибірці даних. Статистичний тест або статистичний критерій – строгий математичний прийом, за яким приймається або відкидається статистична гіпотеза.

Статистичні гіпотези перевіряються на підставі вибірових дослідів або спостережень, тому ймовірність тієї чи іншої гіпотези ніколи не може бути прийнята однозначно рівною 0 або 1. Замість цього використовуються допустимі рівні ймовірності: α та $1-\alpha$ (або β). У практиці статистичного аналізу найчастіше мають справу з двома конкуруючими гіпотезами: *нульовою гіпотезою*, що позначається H_0 та *альтернативною гіпотезою*, що позначається H_a або H_1 .

Нульова гіпотеза – це твердження, що підлягає перевірці. Альтернативна гіпотеза – це протилежне твердження, яке ми намагаємося спростувати у ході статистичного аналізу. Залежно від величини ймовірності альтернативна гіпотеза може бути прийнята або відкинута. Нульова гіпотеза може бути відкинута, але не прийнята! У разі високого рівня ймовірності правдивості нульової гіпотези кажуть, що гіпотеза не може бути відкинута.

Нехай нульова гіпотеза полягає у тому, що випадкова величина належить розподілу 1 на рисунку 1.2.1. Це записується у такий спосіб: $H_0 : X_i (1)$. Конкуруючої гіпотезою в цьому випадку може бути $H_1: X_i (2)$. Випадкова величина X_i належить розподілу 2 на рис. 1.2.1.

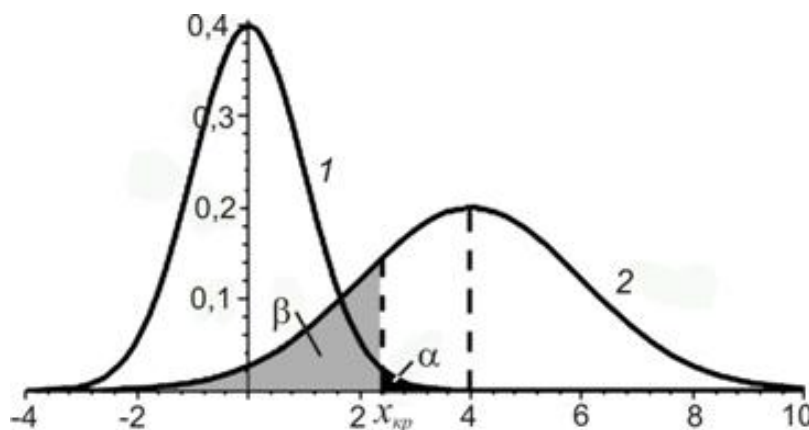


Рис. 1.2.1. Види гіпотез: 1 – нульова; 2 – конкуруюча

Гіпотези бувають прості та складні. Простою називають гіпотезу, яка містить тільки одне припущення. У наведеному вище прикладі обидві гіпотези прості. Складною називають гіпотезу, яка складається з кінцевого або нескінченного числа нульових гіпотез (припущень).

Вибір гіпотези. Щоб з'ясувати, чи відповідає розподіл у генеральній сукупності заданого теоретичного розподілу, необхідно сформулювати досліджувану гіпотезу. В цьому випадку можуть випробовуватися дві гіпотези: 1) $H: \chi^2 = 0$; 2) $H: \chi^2 > 0$.

Для вибору досліджуваної гіпотези необхідно враховувати сліdstва можливих помилок. Обчислюючи теоретичний розподіл, дослідник намагається встановити закономірність даного явища, тобто довести, що досліджуваний розподіл підпорядковується запропонованому теоретичному закону. Тому важливо не зробити помилкового відкриття: чи не приписувати розподілу даний теоретичний закон, якщо насправді воно відповідає іншому закону. Такий висновок відповідає помилковому відхиленню гіпотези $\chi^2 > 0$, якщо вона вірна, тобто коли теоретичний розподіл відрізняється від розподілу генеральної сукупності. Альтернативною повинна бути гіпотеза $\chi^2 = 0$, яка є складовою, і для її перевірки необхідно було б виділити просту гіпотезу, що має практично незначну різницю між розподілами. Однак існує багато теоретичних розподілів, близьких до вибраного. Вказати мінімальне значення χ^2 , яке відповідає різниці між обраним теоретичним та іншим, найближчим до нього, неможливо. Тому у цьому випадку доводиться відступати від загальних правил вибору гіпотез. Використовувана гіпотеза буде нульовою через відсутність різниці між теоретичним і генеральним розподілом, тобто: $H: \chi^2 = 0$. У зв'язку з цим слід вибрати досить великий рівень значущості, щоб зменшити ймовірність, що має більш серйозні наслідки. З огляду на те, що альтернативна гіпотеза $\chi^2 > 0$, проводиться правобічна перевірка.

Якщо вірна альтернативна гіпотеза, то статистична характеристика підпорядковується *нецентральному розподілу* χ^2 . Користуючись цим розподілом, можна обчислити потужність критерію χ^2 . Однак з огляду на особливості вибору гіпотези такі розрахунки потужності у цьому випадку мають невелике практичне значення і не будуть розглядатися.

Методика перевірки статистичних гіпотез. Нехай задана випадкова вибірка $x^m = (x_1, \dots, x_m)$ – послідовність m об'єктів з множини X . Передбачається, що у множині X існує деяка невідома імовірнісна міра P .

Методика полягає в наступному:

1. Формулюється нульова гіпотеза H_0 про розподіл ймовірностей на множині. Гіпотеза формулюється, виходячи з вимог прикладної задачі. Найчастіше розглядаються дві гіпотези – основна або нульова (H_0) і альтернативна (H_1). Іноді альтернатива формулюється у явному вигляді, тоді передбачається, що H_1 означає «не H_0 ». Іноді розглядається відразу кілька альтернатив. У математичній статистиці добре вивчено кілька десятків, які «найбільш часто зустрічаються» типів гіпотез, і відомі ще сотні спеціальних варіантів та різновидів. Приклади наводяться нижче.

2. Задається деяка статистика (функція вибірки) $T: X^m \rightarrow R$, для якої в умовах справедливості гіпотези H_0 виводиться функція розподілу $F(T)$ і / або щільність розподілу $P(T)$. Питання про те, яку статистику треба взяти для перевірки тієї чи іншої гіпотези, часто не має однозначної відповіді. Є цілий ряд вимог, яким повинна задовольняти «добра» статистика T . Висновок функції розподілу $F(T)$ при заданих H_0 і T є суворим математичним завданням, яке вирішується методами теорії ймовірностей. У довідниках наводяться готові формули для $F(T)$, а у статистичних пакетах є готові обчислювальні процедури.

3. Фіксується *рівень значимості* – допустима для даного завдання ймовірність *помилки першого роду*, тобто того, що гіпотеза насправді вірна, але буде відкинута процедурою перевірки. Це повинно бути досить мале число $\alpha \in [0, 1]$. На практиці часто вважають $\alpha = 0,05$.

4. На безлічі допустимих значень статистики T виділяється критична множина Ω_α найменш ймовірних значень статистики T , таке, що $P\{T \in \Omega_\alpha | H_0\} = \alpha$. Обчислення меж критичного як функції від рівня значущості α є суворим математичним завданням, яке у більшості практичних випадків має готове просте рішення.

5. Власне *статистичний тест (статистичний критерій)* полягає в перевірці умови:

- якщо $T(X^m) \in \Omega_\alpha$, то робиться висновок «дані суперечать нульовій гіпотезі при рівні значущості α ». Гіпотеза відкидається.
- якщо $T(X^m) \notin \Omega_\alpha$, то робиться висновок «дані не суперечать нульовій гіпотезі при рівні значущості α ». Гіпотеза приймається.

Отже, *статистичний критерій* визначається статистикою T та критичною множиною Ω_α , яка залежить від рівня значущості.

Зауваження. Якщо дані не суперечать нульовій гіпотезі, це ще не означає, що гіпотеза вірна. Для цього є дві причини:

- У міру збільшення довжини вибірки нульова гіпотеза може спочатку прийматися, але потім виявляються більш тонкі невідповідності даних гіпотезі й вона буде відкинута. Тобто багато що залежить від обсягу даних. Якщо даних не вистачає, можна прийняти навіть саму неправдоподібну гіпотезу.
- Обрана статистика може відображати не всю інформацію, що міститься у гіпотезі H_0 . У такому випадку збільшується ймовірність помилки другого роду – нульова гіпотеза може бути прийнята, хоча насправді вона невірна. Припустимо, наприклад, що $H_0 =$ «Розподілу нормальному», $T(X^m) =$ «Коефіцієнту асиметрії». Тоді вибірка з будь-яким симетричним розподілом буде визнана нормальною. Щоб уникати таких помилок, слід користуватися більш потужними критеріями.

Критерії перевірки статистичних гіпотез, типи та їх властивості

Ряди розподілу є найповнішою характеристикою статистичних сукупностей. Кожному ряду розподілу притаманна певна закономірність, яка виражається залежністю між варіантами та частотами. Закономірності розподілу можна зобразити графічно. Їх можна також висловити аналітично у вигляді закону розподілу, який встановлює відношення варіант ймовірностей. Використовуючи закон розподілу, можна глибше проаналізувати досліджуване явище, спрогнозувати розподіл і т. п. Крім того, необхідною умовою застосування багатьох критеріїв є нормальність розподілу. У зв'язку з цим нерідко потрібно перевіряти відповідність розподілу цієї величини нормальному закону.

Надалі ряд розподілу буде називатися *емпіричним*, а відповідний йому закон – *теоретичним розподілом*. Існує багато законів розподілу: нормальний, логарифмічно-нормальний, Пуассона та ін. Вибір теоретичного закону розподілу повинен базуватися на розумінні механізму формування досліджуваного явища. Його можуть також полегшити графіки емпіричних розподілів.

Частки емпіричного розподілу відрізняються від ймовірностей теоретичного розподілу. Це може бути викликано, наприклад, помилковим вибором теоретичного розподілу (у генеральній сукупності відповідає іншому закону) або випадковими похибками, які обумовлені відхиленнями вибіркового часток від відповідних їм ймовірностей. Тому виникає завдання: за допомогою *критеріїв узгодження* перевірити за даними вибіркового емпіричного розподілу гіпотезу про закон розподілу у генеральній сукупності.

У *логарифмічно-нормальному розподілі* розбіжності між теоретичними та емпіричними частотами значно менші, ніж при нормальному розподілі.

Для перевірки гіпотези про відповідність обраних законів розподілу і розподілу у генеральній сукупності в більшості критеріїв використовуються відхилення емпіричних частот від теоретичних. Очевидно, що чим більше ці відхилення, тим гірше теоретичний розподіл описує емпіричний. Статистичні характеристики більшості критеріїв узгодження є деякими функціями цих відхилень.

Критерій узгодження χ^2 – статистична характеристика

З усіх критеріїв узгодження критерій χ^2 застосовується частіше інших. Його статистичною характеристикою є сума часток від відношення квадрата різниці між емпіричними і теоретичними частотами на теоретичні частки:

$$\chi^2 = \sum_{i=1}^m \frac{(f_i - f'_i)^2}{f'_i}, \quad (1.2.1)$$

де: f_i частки емпіричного розподілу; f'_i частки теоретичного розподілу.

При розрахунку χ^2 частки можна замінити частотами:

$$\chi^2 = n \sum_{i=1}^m \frac{(\omega_i - p_i)^2}{p_i}. \quad (1.2.2)$$

Якщо розподіл у генеральній сукупності відповідає обраному теоретичному закону розподілу, то у генеральній сукупності $\chi^2 = 0$. Чим більше відхилення між цими розподілами, тим більше величина χ^2 . Отже, χ^2 є ступенем відхилення запропонованого теоретичного розподілу від розподілу у генеральній сукупності, а значення χ^2 , обчислене за даними вибірки, – оцінкою цього відхилення. За умови вірно обраного теоретичного розподілу (у генеральній сукупності $\chi^2 = 0$), розподіл статистичної характеристики гіпотези при $n \rightarrow \infty$ зводиться до розподілу χ^2 з $m-l$ ступенями свободи. Критерій можна застосовувати тільки за умови, що усі $f' > 5$. У іншому випадку розподіл статистичної характеристики буде помітно відхилятися від розподілу. Під час визначення кількості ступенів свободи m – кількість груп у емпіричному розподілі, а l – кількість параметрів, які розраховуються за допомогою емпіричного розподілу, i – число незалежних лінійних зв'язків, яким підкоряються емпіричні частки.

Типи статистичних критеріїв

Залежно від конкретної *нульової гіпотези* статистичні критерії поділяються на групи, перераховані нижче за розділами. Поряд з нульовою гіпотезою, яка приймається або відкидається за результатом аналізу вибірки, статистичні критерії можуть спиратися на додаткові припущення, які апріорі припускаються виконаними.

- *Параметричні критерії* передбачають, що вибірка породжена розподілом із заданого параметричного сімейства. Зокрема, існує багато критеріїв, призначених для аналізу вибірок з нормального розподілу. Перевага цих критеріїв у тому, що вони більш потужні. Якщо вибірка дійсно задовольняє додатковим припущенням, то параметричні критерії дають більш точні результати. Однак якщо вибірка їм не задовольняє, то ймовірність помилок (як I, так і II роду) може різко зрости. Перш ніж застосовувати такі критерії, необхідно переконатися, що вибірка задовольняє додатковим припущенням. Гіпотези про вид розподілу перевіряються за допомогою критеріїв згоди.

- *Непараметричні критерії* не спираються на додаткові припущення про розподіл. Зокрема, до цього типу критеріїв відноситься більшість рангових критеріїв.

Критерії згоди перевіряють, чи відповідає задана вибірка з фіксованим розподілом із заданих параметричним сімейством розподілів, або з іншою вибіркою – критерій хі-квадрат і т. п.

Критерії зсуву – це спеціальний випадок двохвибіркових критеріїв згоди. Перевіряється гіпотеза зсуву, згідно з якою розподіли двох вибірок мають однакову форму та відрізняються тільки зрушенням на константу.

Критерії нормальності – це виділений окремий випадок критеріїв згоди. Нормально розподілені величини часто зустрічаються у прикладних задачах, що обумовлено дією закону великих чисел. Якщо про вибірки заздалегідь відомо, що вони підкоряються нормальному розподілу, то до них стає можливо застосовувати більш потужні параметричні критерії. Перевірка нормальності часто виконується на першому кроці аналізу вибірки, щоб вирішити, використовувати далі параметричні методи або непараметричні.

Критерії однорідності призначені для перевірки нульової гіпотези про те, що дві вибірки (або кілька) взяті з одного розподілу, або їх розподіли мають однакові значення математичного очікування, дисперсії, або інші параметри.

Критерії симетричності дозволяють перевірити симетричність розподілу.

Критерії тренду, стаціонарності та випадковості призначені для перевірки нульової гіпотези про відсутність залежності між вибірковими даними та номером спостереження у вибірці. Вони часто застосовуються в аналізі часових рядів, зокрема, при аналізі регресійних залишків.

Властивості статистичних критеріїв

Потужність критерію $1 - \beta(H) = P\{T \in \Omega_\alpha | H\}$: – ймовірність відхилити гіпотезу H_0 , якщо насправді вірна альтернативна гіпотеза H_1 . Потужність критерію є числовою функцією від альтернативної гіпотези H_1 .

Незміщений критерій $1 - \beta(H) \geq \alpha$: для всіх альтернатив H або, що те ж саме $P\{T \in \Omega_\alpha | H\} \geq P\{T \in \Omega_\alpha | H_0\}$, для всіх альтернатив.

Рівномірно більш потужний критерій – критерій з потужністю $1 - \beta(H)$ є рівномірно більш потужним, ніж критерій з потужністю $1 - \beta'(H)$, якщо виконуються дві умови:

1. $\beta(H_0) = \beta'(H_0)$;

2. $\beta(H_1) \leq \beta'(H_1)$ для усіх розглянутих альтернатив $H_1 \neq H_0$, причому хоча б для однієї альтернативи нерівність суворіша.

Статистичний висновок як основа узагальнення даних

Статистичні висновки використовуються для узагальнення даних з вибірки щодо всієї генеральної сукупності. Випадкові помилки, характерні для вибіркового дослідження, можуть привести до того, що вибірка не буде досить точною моделлю генеральної сукупності. Насправді вибірка ніколи не є моделлю генеральної сукупності на всі 100 %, а лише її більш-менш спотвореним варіантом. Для того, щоб оцінити такі спотворення та зробити більш точні висновки про генеральну сукупність і використовуються статистичні висновки. Перш за все, вони дозволяють оцінити ймовірність того, що виявлені у вибірці: взаємозв'язки, відмінності і т. п. – характерні виключно для вибірки, але не для генеральної сукупності. Логіка тут наступна: якщо задана ймовірність висока, то приймається рішення, згідно з яким параметри вибірки не характерні для генеральної сукупності та навпаки – якщо задана ймовірність низька, то прийнято вважати, що відповідні параметри вибірки говорять про параметри генеральної сукупності.

Важливо пам'ятати – досягнення 100 % гарантії того, що результати, отримані у дослідженні, характерні щодо генеральної сукупності, можливі лише у тому випадку, коли проведено суцільне дослідження, тобто опитування всіх представників генеральної сукупності. Але це вже не вибіркове дослідження і воно не передбачає використання статистичних висновків. У найзагальнішому вигляді статистичні висновки можна розділити на дві групи: 1) інтервальне оцінювання (побудова інтервалу, у який із заданою вірогідністю має потрапити середнє значення або пропорція генеральної сукупності); 2) перевірка статистичних гіпотез (ймовірнісний висновок про те, що певні параметри

вибіркової сукупності відображають (або ж ні) параметри генеральної сукупності).

Інтервальне оцінювання у багатьох випадках буває необхідно на основі одиничного параметра вибірки (середнє значення або пропорція) – оцінити відповідний параметр генеральної сукупності. Якщо вибірка достатньо велика (100 і більше спостережень), використовуючи властивості кривої нормального розподілу, можна побудувати інтервал, у який із заданою вірогідністю потрапить справжнє значення. Розподіл вибірових середніх має нормальний вигляд. Відповідно ймовірність отримання вибірки із середнім близьким до середнього значення генеральної сукупності досить висока. Але навіть у тих випадках, коли середнє для вибірки буде досить сильно відрізнятися від середнього генеральної сукупності, довірчий інтервал у переважній більшості випадків буде включати справжнє значення. І лише у дуже рідкісних випадках будуть отримані вибірки, у яких настільки сильно відрізняються параметри вибірки та генеральної сукупності, що справжнє значення не потрапить у довірчий інтервал.

Довірчі інтервали для середніх значень:

$$c.i. = \bar{X} \pm Z \frac{s}{\sqrt{N-1}}, (1.2.3)$$

де: Z – стандартизоване значення, яке визначається рівнем альфа або p -значенням (ймовірністю того, що істинне значення не потрапить у довірчий інтервал);

s – стандартне відхилення за вибіркою;

N – розмір вибірки.

Очевидно що s і N відомі нам з самого дослідження. У свою чергу, Z визначається за допомогою таблиці стандартизованих значень, які рекомендовано застосовувати в аналізі бізнес-процесів:

Таблиця 1.2.1

Довірчий рівень	Альфа (p)	Z-значення
90%	0,10	±1,65
95%	0,05	±1,96
99%	0,01	±2,58
99,9%	0,001	±3,29

Довірчий рівень говорить про те, з якою ймовірністю істинне значення потрапить у побудований інтервал. Розглянемо приклад. Припустимо з вибірки ($N = 300$) відомо, що середнє значення для місячного заробітку киян становить 4000 грн., а стандартне відхилення – 75 грн. Якщо нас задовольнить ймовірність помилки, то вона дорівнює 5% (альфа – 0,05), то $Z = \pm 1,96$.

Звідси:

$$c.i. = 4000 \pm 1,96 \frac{75}{\sqrt{300-1}} = 4000 \pm 8,5; \quad (1.2.3)$$

Таким чином, справжнє значення для киян з імовірністю 95% має потрапити у інтервал від 3991,5 грн. до 4008,5 грн.

$$c.i. = P_s \pm Z \sqrt{\frac{P_u(1-P_u)}{N}}; \quad (1.2.4)$$

У порівнянні з попередньою у цій формулі залишився розмір вибірки і Z -значення. Останнє визначається також за допомогою наведеної вище таблиці 1.2.1. Інші складові такі: P_s – значення пропорції для вибірки; P_u – значення пропорції щодо генеральної сукупності. Фактично дослідник повинен піти на явне збільшення інтервалу, адже більший інтервал з більшою ймовірністю буде включати шукану справжню величину щодо генеральної сукупності. Таким значенням є 0,5: $0,5(1 - 0,5) = 0,5 * 0,5 = 0,25$.

Тепер розглянемо приклад. Нехай, згідно з результатами передвиборчого опитування, за опозиційну партію готові проголосувати 17% населення ($P_s = 0,17$), розмір вибірки N становить 1200 осіб, а рівень альфа – 0,01, тоді:

$$c.i. = 0,17 \pm 2,58 \sqrt{\frac{0,5(1-0,5)}{1200}} = 0,17 \pm 0,02; \quad (1.2.4)$$

Отже, з імовірністю у 99 % за опозиційну партію проголосує від 15 % до 19 % населення.

Дисперсійний аналіз у дослідженні бізнес-процесів

Дисперсійний аналіз (Analysis of variance, або скорочено ANOVA) призначений для перевірок гіпотез у складних випадках. Він являє собою

групу методів порівняння розмірів варіації, обумовлених різними факторами. Дисперсію використовують не тільки для оцінки варіації, а й при вимірюванні взаємозв'язків, для перевірки статистичних гіпотез і т. п. Для ознак метричної шкали розрахунок дисперсії ведеться за формулою:

$$\sigma^2 = \frac{\sum_1^m (x_j - \bar{x})^2 f_j}{\sum_1^m f_j} = \overline{x^2} - \bar{x}^2. (1.2.5)$$

Як і будь-яка середня, дисперсія має певні математичні властивості:

а) якщо всі значення ознаки x_j зменшити (збільшити) на певну величину, дисперсія не зміниться;

б) якщо всі значення ознаки змінити у k раз, то дисперсія зміниться у k^2 разів;

в) у разі заміни частот частками дисперсія не зміниться.

Для альтернативної ознаки, варіація якої має два взаємовиключних значення – «1» та «0», а розподіл характеризується відповідно двома частками – d_1 та d_0 , дисперсія розраховується як добуток часток $\sigma^2 = d_1 d_0 = d_1(1 - d_1)$.

У таблиці. 1.2.2 наведено розрахунок абсолютних характеристик варіації на прикладі терміну обертання облігацій:

Таблиця 1.2.2

Термін обертання облігацій, міс.	f	x_j	$x_j - \bar{x}$	$ x_j - \bar{x} f$	$(x_j - \bar{x})^2 f$
До 2	15	1	-4,6	69,0	317,40
2–4	13	3	-2,6	33,8	87,88
4–6	29	5	-0,6	17,4	10,44
6–8	22	7	1,4	30,8	43,12
8–10	12	9	3,4	40,8	138,72
10 і більше	9	11	5,4	48,6	262,44
Всього	100	X	X	240,4	860,00

Середній термін обертання облігацій 5,6 міс.; середнє лінійне відхилення становить $= 240,4 : 100 = 2,4$ міс.; дисперсія, яка розрахована за формулою (1.2.5) $= 860 : 100 = 8,6$; середнє квадратичне відхилення $= 2,9$ міс.

Частка облігацій з терміном обігу менше 2 міс. становить $d_1 = 0,15$. Дисперсія частки $= 0,15 \cdot (1 - 0,15) = 0,1275$.

При порівнянні варіацій різних ознак або однієї ознаки у різних сукупностях використовують відносні характеристики варіації. Коефіцієнти варіації розраховуються як відношення абсолютних, іменованих характеристик варіації до центру розподілу і часто виражаються у процентах.

Дисперсія, на відміну від інших характеристик варіації, є адитивною величиною. Тобто у структурованій сукупності, яка поділена на групи за факторною ознакою x , дисперсія результативної ознаки y може бути розкладена на: дисперсію у кожній групі (внутрішньогрупову) та дисперсію між групами (міжгрупову). Загальна дисперсія характеризує варіацію ознаки y за рахунок впливу усіх причин (факторів), міжгрупова – за рахунок фактора x , покладеного в основу групування, а внутрішньогрупові – за рахунок інших факторів, не врахованих у групуванні.

Міжгрупова дисперсія обчислюється за формулою:

$$\delta^2 = \frac{\sum_1^m (\bar{y}_j - \bar{y})^2 f_j}{\sum_1^m f_j}; \quad (1.2.6)$$

де: \bar{y}_j та \bar{y} – відповідно середня j -ї групи та загальна середня варіюючої ознаки y ; f_j – частота j -ї групи.

Внутрішньогрупова дисперсія розраховується окремо для кожної j -ї групи:

$$\sigma_j^2 = \frac{\sum_1^{f_j} (y_j - \bar{y}_j)^2}{f_j}; \quad (1.2.7)$$

де y – значення ознаки у окремих елементів сукупності.

Для усіх груп у цілому обчислюється *середня з внутрішньогрупових дисперсій*, зважених на частоти відповідних груп:

$$\overline{\sigma_j^2} = \frac{\sum_1^m \sigma_j^2 f_j}{\sum_1^m f_j}; \quad (1.2.8)$$

Взаємозв'язок між трьома дисперсіями отримав назву правила складання дисперсій, відповідно до якого:

$$\sigma^2 = \delta^2 + \overline{\sigma_j^2}; \quad (1.2.9)$$

Загальну дисперсію можна визначити і безпосередньо по формулі:

$$\sigma^2 = \overline{y^2} - \bar{y}^2; (1.2.10)$$

Відношення міжгрупової дисперсії до загальної називається *кореляційним відношенням*, η^2 .

Розрахунок зазначених дисперсій показано за даними таблиці 1.2.3 на прикладі варіації урожайності винограду:

Таблиця 1.2.3

Сорт винограду	Кількість кущів, f_i	Урожайність з одного куща, кг y	Σy	\bar{y}_i
Аліготе	5	2,3; 2,2; 2,7; 2,6; 2,2	12	2,4
Фетяска	8	2,9; 3,3; 2,7; 2,8; 3,2; 3,4; 2,8; 2,9	24	3,0
Рислінг	7	3,6; 4,0; 4,1; 4,4; 3,8; 3,9; 4,2	28	4,0
У цілому	20	X	64	3,2

Варіація урожайності окремо для кожного сорту винограду визначається трьома внутрішньогруповими дисперсіями. Для сорту винограду «Аліготе» внутрішньогрупова дисперсія становить:

$$\sigma_1^2 = \frac{(2,3-2,4)^2 + (2,2-2,4)^2 + (2,7-2,4)^2 + (2,6-2,4)^2 + (2,2-2,4)^2}{5} = 0,044; (1.2.7)$$

$$\text{де } \bar{y}_1 = \frac{\Sigma y}{f_1} = \frac{12}{5} = 2,4.$$

Аналогічно обчислені внутрішньогрупові дисперсії для другої та третьої груп: = 0,06. Середня з групових дисперсій дорівнює:

$$\overline{\sigma^2} = \frac{0,044 \cdot 5 + 0,06 \cdot 8 + 0,06 \cdot 7}{20} = 0,056;$$

Міжгрупова дисперсія становить:

$$\delta^2 = \frac{(2,4-3,2)^2 \cdot 5 + (3,0-3,2)^2 \cdot 8 + (4,0-3,2)^2 \cdot 7}{20} = 0,4;$$

$$\bar{y} = \frac{\Sigma \bar{y}_j f_j}{\Sigma f_j} = \frac{64}{20} = 3,2.$$

Загальну дисперсію урожайності визначимо за правилом складання дисперсій як суму між групової та середньої з групових дисперсій:

$$\sigma^2 = 0,4 + 0,056 = 0,456$$

Безпосереднє обчислення загальної дисперсій за спрощеною формулою:

$$\sigma^2 = \frac{213,92}{20} - 3,2^2 = 10,696 - 10,24 = 0,456;$$

дає такий самий результат.

У практичній діяльності однофакторний аналіз дисперсій здійснюється за допомогою *MS Excel*. Результат програма виводить у вигляді двох таблиць SUMMARY та ANOVA. При цьому вихідні дані формуються за допомогою блоку команд *Пакету аналізу (Data Analysis)* [4, с. 106–108]. Якщо ваші дані представляють більш складну структуру, дисперсійний аналіз може бути перетворений для відповіді на більш складні питання [25, с. 828–834].

Контрольні питання

1. Сформулюйте визначення поняття «статистична гіпотеза». Наведіть приклади.
2. Схарактеризуйте поняття статистичного критерію.
3. Які типові значення рівня статистичної значущості рекомендовано застосовувати в аналізі бізнес-процесів?
4. Значення статистичного висновку у бізнес-аналізі.
5. Які завдання вирішують за допомогою дисперсійного аналізу?
6. Математичні властивості загальної дисперсії.
7. Схарактеризуйте внутришньогрупову та міжгрупову дисперсії.
8. Застосування пакету прикладних програм для проведення однофакторного та двофакторного аналізу.

1.3. Практичний тренінг розділу 1

Тестові завдання щодо самоконтролю

Структури даних.

Класифікація, розподіл та перетворення даних.

1. Багатовимірні дані містять інформацію:
 - А) про дві ознаки для кожного з об'єктів;
 - В) про три або більше ознак для кожного об'єкта;
 - С) тільки про одну ознаку для кожного об'єкта;
 - Д) усі відповіді вірні.
2. Яких типів бувають якісні дані:
 - А) порядкові й комбіновані;
 - В) номінальні й комбіновані;
 - С) номінальні й порядкові;
 - Д) порядкові й рангові.
3. Яких типів бувають кількісні дані:
 - А) дискретні та безперервні;
 - В) дискретні та порядкові;
 - С) рангові та безперервні;
 - Д) безперервні та комбіновані.
4. Який вид шкали використовується для реєстрації самого нижчого рівня вимірювань:
 - А) шкала інтервалів;
 - В) номінальна шкала;
 - С) логарифмічна шкала;
 - Д) шкала пропорцій.
5. Прикладом якої шкали є міри довжини (м, см. і т. п.):
 - А) шкала відносин;
 - В) рангова шкала;
 - С) порядкова шкала;
 - Д) шкала інтервалів.
6. Удосконалення управління зумовлює використання менеджерами результатів статистичного моніторингу ділової активності

підприємств для розробки подальших можливих корпоративних стратегій розвитку:

- А) реструктуризації;
- В) диверсифікації;
- С) вертикальної інтеграції;
- Д) портфельного аналізу.

7. В основу портфельного аналізу покладено такі оцінні критерії:

- А) теперішня вартість очікуваних доходів від володіння цінними паперами (проценти, дивіденди);
- В) рівень ризикованості вкладень;
- С) аналіз портфеля продуктів (товарів і послуг) підприємства;
- Д) усі відповіді вірні.

Випадкові величини і закони розподілу
як основа бізнес-статистики

8. Ймовірність того, що випадкова величина X (дискретна або безперервна) приймає значення, менше x , називається:

- А) функцією розподілу;
- В) величиною розподілу;
- С) законом розподілу;
- Д) гіпотезою розподілу.

9. Відносною частотою випадкової події A називається:

- А) відношення числа n проведених випробувань до загальної кількості nA появ цієї події;
- В) відношення числа nA появ цієї події до загального числа n проведених випробувань;
- С) відношення числа nA появ цієї події до середнього числа n проведених випробувань;
- Д) усі відповіді вірні.

10. Який вид розподілу визначається двома параметрами, n і p :

- А) геометричний розподіл;
- В) рівномірний розподіл;
- С) біномінальний розподіл;
- Д) показовий розподіл.

- 11.** Дві випадкові величини X і Y називаються незалежними, якщо:
- А) незалежні всі пов'язані з ними події;
 - В) незалежним є хоча б одна пов'язана з ним подія;
 - С) жодна пов'язана з ним подія не є незалежною;
 - Д) усі відповіді невірні.
- 12.** Коваріація двох випадкових величин X і Y характеризує:
- А) ступінь їх залежності навколо точки (m_x, m_y) ;
 - В) міру їх розсіювання навколо точки (m_x, m_y) ;
 - С) ступінь їх залежності та міру розсіювання навколо точки (m_x, m_y) ;
 - Д) ні одна з відповідей не є вірною.
- 13.** Безрозмірна величина, що характеризує тільки залежність випадкових величин X і Y , а не розкид, називається:
- А) коефіцієнт детермінації;
 - В) коефіцієнт Спірмена;
 - С) коефіцієнт Чупрова;
 - Д) коефіцієнт кореляції.
- 14.** Коефіцієнт кореляції характеризує:
- А) ступінь лінійної залежності випадкових величин X і Y ;
 - В) ступінь логарифмічної залежності випадкових величин X і Y ;
 - С) ступінь експоненційної залежності випадкових величин X і Y ;
 - Д) ступінь інтегральної залежності випадкових величин X і Y .

Статистичні гіпотези та методика їх перевірки

- 15.** Статистична гіпотеза – це:
- А) певне твердження, яке потрібно підтвердити або відкинути на основі наявних статистичних даних;
 - В) інформація, яку отримують на підставі даних генеральної сукупності;
 - С) ступінь ризику, що відповідає тому чи іншому висновку;
 - Д) ймовірність того, що прийняті рішення будуть помилковими.
- 16.** За прикладним змістом можна виділити такі групи статистичних гіпотез:
- А) щодо типу розподілу досліджуваної випадкової величини;
 - В) щодо числових характеристик досліджуваної генеральної сукупності;

- С) щодо однорідності двох і більше вибірок або характеристик досліджуваних сукупностей;
- Д) щодо ступеню та типу залежності між досліджуваними ознаками.
- 17.** Гіпотезу, яка заперечує нульову, називають:
- А) параметричною;
- В) нульовою;
- С) конкуруючою;
- Д) альтернативною.
- 18.** Прийняти неправильне рішення під час перевірки статистичної гіпотези можна у:
- А) одному випадку;
- В) двох випадках;
- С) трьох випадках;
- Д) трьох і більше випадках.
- 19.** Точки, що відділяють критичну область від області прийняття гіпотез, називають:
- А) критичними;
- В) односторонніми;
- С) двохсторонніми;
- Д) правосторонніми.
- 20.** До завдань перевірки нульової гіпотези не входить:
- А) вибір рівня істотності;
- В) визначення критичної області;
- С) обчислення статистичного критерію;
- Д) оцінка похибки вибірки.
- 21.** Якщо перевіряють гіпотези щодо вибірових середніх за невідомого значення дисперсії генеральної сукупності, застосовують критерій, що відповідає:
- А) розподілу Фішера;
- В) розподілу Стюдента;
- С) нормальному розподілу;
- Д) варіаційному розподілу.

Дисперсійний аналіз у дослідженні бізнес-процесів

- 22.** З англійської мови ANOVA перекладається як:
- А) кореляційний аналіз;
 - В) регресійний аналіз;
 - С) математичний аналіз;
 - Д) дисперсійний аналіз.
- 23.** Виберіть вірну математичну властивість, характерну для дисперсії:
- А) якщо всі значення ознаки змінити у k разів, то дисперсія зміниться у k^2 разів;
 - В) якщо всі значення ознаки збільшити у k разів, то дисперсія зменшиться у k разів;
 - С) якщо всі значення ознаки зменшити у k разів, то дисперсія збільшиться у k разів;
 - Д) жодне з вище згаданих властивостей не характерно щодо дисперсії.
- 24.** Якщо усі значення ознаки x_j зменшити або збільшити на певну величину, то:
- А) при збільшенні ознаки дисперсія збільшиться, при зменшенні – зменшиться;
 - В) дисперсія не зміниться;
 - С) при зменшенні ознаки – дисперсія збільшиться;
 - Д) при збільшенні ознаки – дисперсія зменшиться.
- 25.** У структурованій сукупності, розділеної на групи за ознакою x , дисперсія ознаки y може бути розкладена на:
- А) міжгрупові та зовнішньогрупові;
 - В) зовнішньогрупову та внутрішньогрупову;
 - С) міжгрупову та внутрішньогрупову;
 - Д) міжгрупову та середньогрупову.
- 26.** Щодо усіх груп у цілому обчислюється:
- А) середня з міжгрупових;
 - В) середня з внутрішньогрупових;
 - С) середня з зовнішньогрупових;
 - Д) середня арифметична.

27. Взаємозв'язок між трьома дисперсіями має такий вигляд:

- А) загальногруппова дисперсія = міжгруппова – внутрішньогруппова;
- В) загальногруппова дисперсія = міжгруппова + внутрішньогруппова;
- С) внутрішньогруппова дисперсія = міжгруппова * загальну дисперсію;
- Д) міжгруппова дисперсія = загальна дисперсія : середню з внутрішньогруппових.

28. Ставлення групової дисперсії до загальної називається:

- А) варіаційним відношенням η^2 ;
- В) дисперсійним відношенням η^2 ;
- С) емпіричним відношенням η^2 ;
- Д) кореляційним відношенням η^2 .

Розділ 2

СТАТИСТИЧНА ОЦІНКА ЕФЕКТИВНОСТІ СТАНУ РОЗВИТКУ БІЗНЕСУ В КРАЇНІ

2.1. Непараметричні методи та аналіз закономірностей для якісних даних у бізнес-аналізі

*Сутність і передумови використання непараметричних методів
у бізнес-статистиці. Рангова кореляція.*

Оцінка зв'язків між альтернативними ознаками у бізнес-статистиці.

Якісні дані у бізнес-статистиці та особливості їх аналізу.

*Статистичний критерій χ^2 -квадрат:
характеристика та методи розрахунку.*

*Виявлення прихованих закономірностей
для вирішення бізнес-задач за допомогою Data-Mining*

*Сутність і передумови використання непараметричних методів
у бізнес-статистиці. Рангова кореляція.*

Непараметрична статистика – це самостійна галузь математичної статистики, яка набула популярності у другій половині XIX століття. Саме тоді були розроблені такі класичні методи аналізу даних, що можуть бути виражені лише у номінальній та ранговій шкалах вимірювання, про що йшла мова у темі 1.1. Методологічною основою непараметричної статистики є непараметричні методи дослідження. До них відноситься статистична перевірка гіпотез на основі даних, які не мають характеру нормального розподілу, оскільки передбачають використання частот або рангів.

Загалом непараметричні методи є більш гнучкими і не вимагають точного визначення ситуації. Прикладами можуть бути значення експорту чи імпорту продукції до і після введення митного тарифу чи квот; результати опитувань інвесторів щодо рівня інвестиційної привабливості підприємств до та після проведення реформ; ставлення чоловіків та їх дружин до купівлі-продажу нерухомості за участі маклерів тощо.

Виявлення причинно-наслідкових зв'язків у характеристиках бізнес-процесів та прийняття ефективних управлінських рішень часто потребують використання математичного апарату, який доволі простий та якому притаманна стійкість висновків. Найбільш популярним є коефіцієнт кореляції рангів Спірмена, який було названо на честь Чарльза Едварда Спірмена – англійського психолога, відомого своїми працями у галузі статистики.

Розглянемо розрахунок цього коефіцієнта на прикладі.

Приклад. За даними таблиці 2.1.1 оцінимо тісноту зв'язку між рівнем ефективності економіки і надійністю ділового партнерства для семи країн Східної Європи. Оскільки інформація представлена у формі інтегральних показників (оцінки у балах), необхідно провести ранжування країн. Найменшим значенням інтегрального показника надається ранг 1, найбільшому – ранг 7. Сума квадратів відхилень рангів становить:

$$\sum_1^n d_j^2 = 10,$$

а коефіцієнт рангової кореляції:

$$\rho = 1 - \frac{6 \sum_1^n d_j^2}{n(n^2 - 1)} \quad (2.1.1)$$

$$\rho = 1 - \frac{6 * 10}{7(7^2 - 1)} = 1 - \frac{60}{336} = 0,82$$

Таблиця 2.1.1

Країна	Інтегральні показники		Ранги показників		Відхилення рангів d_j	d_j^2
	ефективність економіки ($max = 10$)	надійність ділового партнерства ($max = 100$)	R_x	R_y		
A	5,9	54,9	6	7	-1	1
B	7,1	54,8	7	6	1	1
C	4,2	45,3	4	5	-1	1
D	3,4	36,9	3	4	-1	1
K	4,9	35,8	5	3	2	4
M	2,7	26,4	1	2	-1	1
P	2,9	24,8	2	1	1	1
Всього	X	X	X	X	0	10

Значення коефіцієнта рангової кореляції свідчить про наявність прямого і досить помітного зв'язку між зазначеними параметрами ризику іноземного інвестування економіки. Табличне критичне значення коефіцієнта рангової корекції для $\alpha = 0,05$ та $n = 7$ становить $\rho_{0,95}(7) = 0,71$, що значно менше фактичного. Отже, істотність зв'язку доведена з ймовірністю 0,95.

Оцінка зв'язків між альтернативними ознаками
у бізнес-статистиці

Аналіз взаємозв'язків між атрибутивними ознаками проводиться на основі таблиць взаємної спряженості (взаємозалежності), що описують комбінаційні розподіли сукупностей за двома ознаками – факторному x та результативному y . При наявності стохастичного зв'язку умовні розподіли змінюються від групи до групи. Оцінка тісноти стохастичного зв'язку ґрунтується на відхиленнях частот (частостей) умовних розподілів від безумовного, тобто на відхиленнях фактичних частот f_{ij} від теоретичних F_{ij} , пропорційних підсумковим частотам безумовного розподілу:

$$F_{ij} = \frac{f_{i0}f_{0j}}{n}, \quad (2.1.2)$$

де: f_{i0} – підсумкові частоти за ознакою x ;

f_{0j} – підсумкові частоти за ознакою y ,

n – обсяг сукупності.

Очевидно, що:

$$n = \sum_i^{m_x} f_{i0} = \sum_j^{m_y} f_{0j}; \quad (2.1.3)$$

Абсолютну величину відхилень $(f_{ij} - F_{ij})$ характеризує квадратична спряженість Пірсона χ^2 :

$$\chi^2 = \sum_i \sum_j \frac{(f_{ij} - F_{ij})^2}{F_{ij}} = n \left[\sum_i \sum_j \frac{f_{ij}^2}{f_{i0}f_{0j}} - 1 \right]; \quad (2.1.4)$$

Якщо стохастичний зв'язок відсутній, то $\chi^2 = 0$. Для того, щоб зробити висновок про суттєвості зв'язку, необхідно порівняти фактичне значення χ^2 з критичним для заданої ймовірності $1 - \alpha$ та числа ступенів

свободи $k = (m_x - 1)(m_y - 1)$, де: m_x та m_y – відповідно кількість груп за ознаками x та y .

Відносною мірою тісноти стохастичного зв'язку слугують коефіцієнти взаємної спряженості C , які за змістом ідентичні коефіцієнтам кореляції. Якщо $m_x = m_y$, використовують коефіцієнт спряженості Чупрова:

$$C = \sqrt{\frac{\chi^2}{n\sqrt{(m_x-1)(m_y-1)}}}; \quad (2.1.5);$$

якщо $m_x \neq m_y$, перевагу віддають коефіцієнту спряженості Крамера:

$$C = \sqrt{\frac{\chi^2}{n(m_{\min}-1)}}, \quad (2.1.6)$$

де: m_{\min} – мінімальна кількість груп за ознакою x або y .

Значення коефіцієнта C коливаються в межах від 0 до 1.

Приклад. У таблиці 2.1.2 наведено комбінаційний розподіл респондентів як потенційних інвесторів за віком та схильністю до ризику. До групи ризикових віднесено респондентів, які мають намір придбати цінні папери, незважаючи на ризик, обережні не ризикують без гарантій, неризикові уникають ризику взагалі. Концентрація частот навколо діагоналі з верхнього лівого кута у правий нижній свідчить про наявність стохастичного зв'язку.

Таблиця 2.1.2

Вік, років	Тип інвестора			Разом
	ризиковий	обережний	неризиковий	
До 30	24	12	4	40
30–50	20	50	30	100
50 і старше	6	18	36	60
Разом	50	80	70	200

Фактичне значення χ^2 становить:

$$\chi^2 = 200 \left[\frac{24^2}{40 \cdot 50} + \frac{12^2}{40 \cdot 80} + \frac{4^2}{40 \cdot 70} + \frac{20^2}{100 \cdot 50} + \frac{50^2}{100 \cdot 80} + \frac{30^2}{100 \cdot 70} + \frac{6^2}{60 \cdot 50} + \frac{18^2}{60 \cdot 80} + \frac{36^2}{60 \cdot 70} - 1 \right] = 49,6$$

що значно вище критичного $\chi^2_{0,95}(4) = 9,49$. Отже, істотність зв'язку між віком респондентів і схильністю їх до ризику доведена з ймовірністю 0,95.

Оскільки $m_x = m_y = 3$ для оцінки щільності зв'язку використовуємо коефіцієнт взаємної спряженості Чупрова:

$$C = \sqrt{\frac{49,5}{200 \sqrt{2 \cdot 2}}} = 0,352, \quad (2.1.5)$$

тобто зв'язок між ознаками помірний.

У разі, коли $m_x = m_y = 2$, розрахунок коефіцієнта взаємної спряженості спрощується:

$$C = \sqrt{\frac{\chi^2}{n}} = \frac{f_{11}f_{22} - f_{12}f_{21}}{\sqrt{f_{01}f_{02}f_{10}f_{20}}}. \quad (2.1.7)$$

У статистичній літературі коефіцієнт C для 4-клітинної таблиці називають *коефіцієнтом контингенції (асоціації)*. Очевидно, що

$$\chi^2 = nC^2. \quad (2.1.8)$$

Для аналізу таблиць такого типу використовують також відношення перехресних добутоків, або *відношення шансів*:

$$W = \frac{f_{11}f_{22}}{f_{12}f_{21}}. \quad (2.1.9)$$

Відношення шансів характеризує міру відносного ризику (виграшу), пов'язаного з фактором x на результат y .

Приклад. За даними таблиці 2.1.3 оцінимо щільність зв'язку між сприйняттям реклами та придбанням рекламованого товару, а також результативність реклами:

Таблиця 2.1.3

Сприйняття реклами	Кількість респондентів		Разом
	придбали товар	не придбали товар	
Запам'ятали рекламу	9	31	40
Не запам'ятали рекламу	2	38	40
Разом	11	69	80

Коефіцієнт контингенції свідчить про наявність стохастичного зв'язку

$$C = \frac{9 \cdot 38 - 2 \cdot 31}{\sqrt{11 \cdot 69 \cdot 40 \cdot 40}} = 0,254. \quad (2.1.7)$$

Фактичне значення $\chi^2 = nC^2 = 80 \cdot 0,254^2 = 5,16$, що вище критичної позначки $\chi^2_{0,95}(1) = 3,84$. Отже, істотність зв'язку доведена.

Відношення шансів становить: $W = \frac{9 \cdot 38}{2 \cdot 31} = 5,5$, (2.1.9), тобто шанси

реалізувати рекламований товар у 5,5 раза вище у порівнянні з нерекламованим.

Методи аналізу таблиць взаємної спряженості можна використовувати і для кількісних ознак, проте слід зазначити, що міри щільності кореляційної зв'язку – коефіцієнт детермінації та кореляційне відношення – більш потужні.

Якісні дані у бізнес-статистиці та особливості їх аналізу

Якісні дані, як і кількісні, описують властивості елементів, процесів чи явищ. Для обробки масивів якісних даних часто використовують статистичні, комбінаторні та евристичні показники. Перелік показників, які використовують для аналітичного оцінювання якісних даних подано у таблиці 2.1.4 [4, с. 142].

Таблиця 2.1.4

Статистичні показники аналізу якісних даних у бізнесі [4, с. 142]

Параметри для визначення	Показники	Характеристика	Приклади використання
Середнього значення сукупності	Мода та середній нормований бал	Типовий рівень варіативної ознаки	Модальне значення освіти працівників
Варіації ознак	Дисперсія альтернативної ознаки	Коливання, мінливість значень ознаки	Дисперсія якісної продукції (контроль якості продукції)
Концентрації, диференціації	Коефіцієнт локалізації, коефіцієнт концентрації	Ступінь концентрації ознак у сукупності загалом, в окремих групах	Коефіцієнт концентрації власного капіталу
Аналітичні оцінки взаємозв'язків між ознаками	Коефіцієнти взаємного спряження, коефіцієнт рангової кореляції	Щільність зв'язку між ознаками	Вплив віку менеджерів на прийняття рішень на основі інтуїції

Очевидно, що перелік показників, які обчислюють у бізнес-аналізі даних, є дещо обмежений. Це зумовлено статистичною природою атрибутивних ознак. Статистичні висновки щодо якісних даних у випадку, якщо результати спостереження подано не числами, а категоріями, роблять на основі перевірки гіпотез.

Статистичний критерій χ^2 -квадрат:
характеристика та методи розрахунку

Критерій χ^2 -квадрат використовують для перевірки гіпотез про якісні дані, представлені не числовими, а категоріальними змінними. Тут прийнято оперувати підрахунком частоти (оскільки ранжування або арифметичні дії виконувати неможливо). Критерій (тест) « χ^2 -квадрат» заснований на частотах, які представляють собою кількість об'єктів вибірки, що потрапляють в ту чи іншу категорію. Суть показника χ^2 -квадрат (χ^2): він вимірює різницю між (експериментальними) частотами f_E та очікуваними (теоретичними) частотами f_T . Саме він розраховується як сума квадратів різниці цих частот, виражених у частках теоретичної частоти. Використання такого статистичного підходу можливо у різних обставинах. Розглянемо найбільш поширені.

Комбінація: нинішні й минулі події
(критерій « χ^2 -квадрат» відповідності)

Даний спосіб широко застосовується у тих випадках, коли потрібно визначити, чи є наш нинішній досвід (виражений у частотах або відсотках) типовим по відношенню до минулого досвіду (набір так званих опорних величин). Таку ситуацію можна умовно позначити фразою «Це було недавно, а то було давно. Між ними є відповідність?». Тест « χ^2 -квадрат» щодо відповідності відсотків використовується для перевірки гіпотези про те, що комбінація спостережуваних частот або відсотків (що характеризують одну якісну змінну) побудована на даних з деякою генеральної сукупності з уже відомими значеннями відсотків (опорними величинами). Можна сформулювати висловлені міркування і по-іншому: ті результати, які ми спостерігаємо зараз (фактичні дані, тобто наш нинішній досвід), за характером аналогічні минулим даним (опорним величинам). Це пояснюється тим, що і ті, й інші ставляться до

однієї і тієї ж генеральної сукупності, але витягувалися у різний час (зараз і колись давно). Очікуване значення частоти для кожної категорії розраховується як твір заданого опорного значення відсотка у генеральній сукупності на розмір вибірки n . На підставі наявних знань про те, що спостерігаються очікувані частоти аналізованої події, визначається власне показник χ^2 -квадрат. Розрахункове значення χ^2 -квадрат потім порівнюють з критичним (табличним) показником для відповідного числа ступенів свободи, який визначається як кількість категорій мінус одиниця. Якщо виявляється справедливим нерівність $\chi^2_{\text{расч}} > \chi^2_{\text{крит}}$, то із заданою вірогідністю (або рівнем значущості) можна стверджувати, що спостерігаємі частоти (наш досвід) значимо відрізняються від тих, які очікуються, виходячи з відомих нам опорних значень відсотків (частот). Отже, обґрунтовано можна робити висновок про те, що спостерігаємі вибіркові відсотки значимо відрізняються від заданих опорних значень. Якщо маємо співвідношення $\chi^2_{\text{расч}} < \chi^2_{\text{крит}}$, то спостережувані значення незначно відрізняються від опорних показників і, отже, наші фактичні результати не мають значущих відмінностей від заданих опорних значень. При виконанні такого аналізу прийнято дотримуватися наступного емпіричного правила: очікувані частоти у кожній категорії повинні бути, принаймні, не менше п'яти (оскільки критерій χ^2 -квадрат залишається приблизною, а не точною оцінкою). Аналіз критерію відповідності відсотків (частот) зручно виконувати, дотримуючись наступної схеми.

1. Є табличні дані частот щодо кожної категорії однієї якісної змінної. Обговорюються такі гіпотези:
 - а) частоти (відсотки) нинішнього досвіду рівні набору відомих, фіксованих опорних величин (з минулого досвіду);
 - б) частоти (відсотки) нинішнього досвіду не рівні набору опорних величин (даних минулого досвіду).
2. Очікувані частоти обчислюються так: потрібно для кожної категорії помножити відоме значення її частки у загальній кількості (генеральній сукупності) на розмір вибірки n . При цьому передбачається, що:

а) набір даних представляє собою випадкову вибірку з даної генеральної сукупності;

б) очікується наявність, принаймні, п'яти об'єктів у кожній з категорій;

3. Проводиться аналіз « χ^2 -квадрат».

4. Інтерпретація результату тесту « χ^2 -квадрат»: наявність значущої зв'язку відзначається тоді, коли розрахункове значення « χ^2 -квадрат» більше табличного або критичного (тобто: $\chi^2_{\text{расч}} > \chi^2_{\text{крит}}$), у іншому випадку значущої зв'язку немає. Тепер приступимо до конкретного аналізу критерію відповідності частот. Розглянемо наступний приклад. Серед студентів економічного факультету, які здавали на першому курсі у літню сесію іспит з математики, було проведено опитування з метою з'ясування того, які чинники впливають на отримання незадовільної оцінки. Число опитаних студентів становило 50 осіб. Найбільш часто згаданими причинами були наступні:

1. Сам винен, треба було краще займатися.
2. Строгий викладач.
3. На жаль, не вдалося списати.
4. Позначився вплив фатальних, принаймні (дістався квиток № 13, зустрів чорного кота, забув надіти «щасливий» светр та ін.).

Ці відповіді можна умовно розділити на наступні категорії:

1. Сам винен.
2. Дуже строгий викладач.
3. Шпаргалки.
4. Прикмети.

У таблиці 2.1.5 наведено дані про причини отримання «незадовільної оцінки» з математики за минулу сесію, а також вказані значення опорних величин, взяті з екзаменаційних відомостей з цього предмету за минулі роки (за тими ж категоріями). Як бачимо з даних таблиці 2.1.5, за кількісними показниками усі аналізовані причини формально відрізняються від опорних значень. Однак ця відмінність виявляється далеко нерівноцінно. Так, можна визнати, що у категорії самооцінки «Сам винен» фактичні дані відрізняються від відповідних опорних величин відносно слабо (наприклад, 57 % у порівнянні з 59 % для

минулих сесій). У той же час за іншими категоріями відносна відмінність виглядає більш помітно. Особливо впадає в очі невідповідність з позиції «Шпаргалки».

Таблиця 2.1.5

**Підсумкові дані про причини отримання незадовільної оцінки
з математики за дану сесію та сесії минулих років**

Причина	Спостережувані дані (за минулу сесію)		Опорні значення, % (очікувані дані)
	частота	відсоток від загального числа, %	
Сам винен	28	57	59
Дуже строгий викладач	10	19	14
Шпаргалки	7	14	20
Прикмети	5	10	7
Усього:	50	100	100

Питання полягає у тому, значима чи ця різниця? Інакше кажучи, чи можуть отримані за підсумками минулої сесії «незадовільні оцінки» розглядатися як результат вилучення випадкової вибірки з генеральної сукупності, у якій відсотки «незадовільної оцінки» відповідають опорним величинам? Або ще по-іншому: чи достатньо велика спостерігається різниця, щоб її не можна було пояснити тільки випадковістю? Тест « χ^2 -квадрат» відповідності відсотків дозволить дати відповідь на це питання. Стверджувальний висновок отримаємо за умови, коли виявиться справедливим співвідношення $\chi^2_{\text{расч}} > \chi^2_{\text{крит}}$. Його потрібно буде тлумачити наступним чином: результати нинішньої сесії та результати минулої сесії відрізняються між собою принципово, оскільки відмінність між ними не носить випадкового характеру. Якщо виявиться справедливим нерівність $\chi^2_{\text{расч}} < \chi^2_{\text{крит}}$, то із заданою вірогідністю можна буде говорити про незначимі відмінності між аналізованими результатами. У таблиці 2.1.6 вкажемо частотні величини для обох інформаційних позицій – поточні дані «Спостереження» та відомості за минулі роки «Очікування». Розрахунок частот для графі «Очікування» (тобто очікувані частоти) проведемо шляхом множення значень опорних величин відсотків (59 %, 14 %, 20 % та 7 %) на розмір вибірки ($n = 50$).

У результаті отримаємо такі значення частот: $0,59 \times 50 = 29,5$; $0,14 \times 50 = 7,0$ та т. п. Зауважимо, що у підсумковому рядку для обох колонок загальна сума частот однакова та дорівнює 50.

Таблиця 2.1.6

**Спостережувані та очікувані дані (частоти)
щодо причин незадовільних оцінок**

Причина	«Спостереження»	«Очікування»
Сам винен	28,0	29,5
Дуже строгий викладач	10,0	7,0
Шпаргалки	7,0	10,0
Прикмети	5,0	3,5
Усього	50,0	50,0

Ці дані і будемо використовувати для вирішення питання про значення відповідності (або невідповідності) фактичних та очікуваних результатів.

Коефіцієнти взаємної спряженості

На основі « χ^2 -квадрата» прийнято також оцінювати показники ступеня тісноти зв'язку – коефіцієнти взаємної спряженості К. Пірсона та А. Чупрова.

Коефіцієнт Пірсона розраховується за формулою:

$$K_p = \sqrt{\frac{\chi^2}{n + \chi^2}}; \quad (2.1.10)$$

де: χ^2 – розрахункове значення « χ^2 -квадрата», n – загальне число спостережень (обсяг вибірки).

Коефіцієнт Чупрова дозволяє врахувати число груп з кожної ознаки і визначається наступним чином:

$$K_{\text{ч}} = \sqrt{\frac{\chi^2}{n \sqrt{(k_1 - 1)(k_2 - 1)}}}; \quad (2.1.11)$$

де: k_1 та k_2 – відповідно число значень (груп) для першої та другої ознак або, по-іншому, число рядків та стовпців у таблиці;

n – загальне число спостережень (обсяг вибірки). Спробуємо виконати такі розрахунки для нашого прикладу.

$$K_{\pi} = \sqrt{\frac{\chi^2}{n+\chi^2}} = \sqrt{\frac{2,905}{50+2,905}} = 0,234; (2.1.10)$$

$$K_{\pi} = \sqrt{\frac{\chi^2}{n\sqrt{(k_1-1)(k_2-1)}}} = \sqrt{\frac{2,905}{50\sqrt{(4-1)(2-1)}}} = 0,184; (2.1.11)$$

Розрахунок обох коефіцієнтів дає вельми малі величини, що свідчить про відсутність зв'язку між досліджуваними характеристиками. Це ж підтверджують і оцінки з таблиці Чеддока: розраховані коефіцієнти з модулю менш, ніж 0,3, говорять про відсутність кореляційного зв'язку. Інакше кажучи, використання цих коефіцієнтів підтверджує раніше висловлене міркування: досліджувана ситуація за своїми параметрами відповідає опорним (очікуваним) показникам і тому не вимагає введення будь-яких коригувань.

*Перевірка наявності взаємозв'язку між двома якісними змінними
(критерій «хі-квадрат» незалежності)*

Можливі ситуації, коли є дві якісні змінні, тобто набір експериментальних даних являє собою двовимірні якісні дані. Після вивчення кожної з них окремо за допомогою аналізу частот (або відсотків) може виникнути питання про наявність зв'язку між ними. Вважається, що дві якісні змінні є незалежними, якщо знання значення однієї змінної не допомагає передбачити значення іншої. Уявімо собі, що ваша фірма розробила технологію гальванічного покриття нікелем сталевих деталей автомобільного кузова. У середньому відсоток браку, пов'язаного з відшаруванням покриття, становить 3,1 %. Однак коли працює технолог Іванов, розмір браку досягає 11,2 %. У цьому випадку знання значення однієї змінної (ім'я конкретного технолога) допомагає спрогнозувати значення іншої змінної, оскільки 3,1 % та 11,2 % розрізняються між собою. Отже, ці дві змінні не є незалежними. Використання критерію «хі-квадрат» дозволяє вирішити питання про те, що розглядаються якісні сукупності залежними або ж незалежними один від одного. У цьому випадку застосовується так званий критерій «хі-квадрат» незалежності, який встановлює наявність (або відсутність) зв'язку між двома якісними змінними. Для такого аналізу використовується таблиця частот, які можна було б очікувати у тому

випадку, якщо змінні виявилися б незалежними. У загальному випадку критерій «хі-квадрат» незалежності прийнято представляти наступним чином:

1. Є вихідні дані у формі табличного списку частот усіх комбінацій категорій двох якісних змінних. Обговорюються такі гіпотези:

- а) дві змінні не залежать одна від одної;
- б) дві змінні пов'язані, вони не є незалежними одна від одної.

2. Складається таблиця очікуваних (теоретичних) частот. Для їх розрахунку частоту однієї категорії (результат експерименту) слід помножити на частоту іншої категорії (також експериментальний показник) і отриманий добуток поділити на загальний обсяг вибірки n :

$$\text{Очікувана частота } f_{\text{оч(т)}} = \frac{\text{Частота категорії } f_{e1} \times \text{Частота категорії } f_{e2}}{\text{Загальний об'єм вибірки } n}, \quad (2.1.12)$$

або більш компактно, у символічній формі:

$$f_{\text{оч(т)}} = \frac{f_{e1} \times f_{e2}}{n}; \quad (2.1.12)$$

При цьому вважається, що: а) набір даних представляє собою випадкову вибірку з даної генеральної сукупності та б) для кожної комбінації категорій очікувана частота, принаймі, не менш п'яти.

3. Далі проводиться аналіз «хі-квадрат», розрахунок виконується з використанням знайомого виразу:

$$\chi^2 = \sum \frac{(f_e - f_{\pi})^2}{f_{\pi}}; \quad (2.1.13)$$

Ступінь свободи обчислюється таким чином:

$$f = (k_1 - 1) \times (k_2 - 1); \quad (2.1.14)$$

де: k_1 та k_2 – число категорій відповідно для першої та другої змінної.

4. Результат тесту «хі-квадрат» трактується так: наявність значущої зв'язку проявляється тоді, коли розрахункове значення «хі-квадрат» більше критичного (тобто $\chi^2_{\text{расч}} > \chi^2_{\text{крит}}$), у іншому випадку значущої зв'язку немає. Давайте познайомимось з цим видом статистичного аналізу, для чого розглянемо наступний приклад. Кіт Матроскін, зайнявшись молочним бізнесом, вирішив провести маркетингове

дослідження, щоб з'ясувати, який вид молочної продукції вважають за кращим ті чи інші покупці. Для кожної покупки фіксувалися дві якісні змінні – вид продукції та тип покупця. У якості молочної продукції фігурували молоко, сметана та сир. Покупців Матроскін умовно розділив на дві категорії – практичні та імпульсивні. До перших він відніс тих покупців, які йдуть на ринок уже з чітко сформульованим наміром щодо того, що купити і скільки саме. Другу ж категорію склали покупці, які рішення приймають на місці, безпосередньо перед тим, як зробити покупку. Отримані дані статистичного опитування Матроскін представив у табличній формі (таблиця 2.1.7), у якій щодо кожного виду молочної продукції вказав кількість скоєних покупок тим чи іншим покупцем, тобто вказав фактичну частоту.

Таблиця 2.1.7

Результати опитування щодо перспектив молочного бізнесу

Вид молочної продукції	Частота переваг типів покупців:	
	практичного	імпульсивного
Молоко	38	15
Сметана	24	31
Сир	18	27

Дослідник має надати висновок за підсумками статистичної перевірки за критерієм « χ^2 -квадрат», тобто сформулювати висновок та пояснити результат з практичної точки зору – визначити, яку ринкову стратегію має обрати кіт Матроскін і, отже, на якого покупця та на який вид молочної продукції йому належить орієнтуватися. Для цього доповнимо таблицю 2.1.7 з вихідними даними «підсумковим» рядком та додатковим «сумарним» стовпцем, заповнимо їх, виконавши нескладні розрахунки і представимо у таблиці 2.1.8.

Таблиця 2.1.8

Доповнені дані за результатами опитування щодо перспектив молочного бізнесу

Вид молочної продукції	Частота переваг типів покупців:		Усього
	практичного	імпульсивного	
Молоко	38	15	53
Сметана	24	31	55
Сир	18	27	45
Усього	80	73	153

Чисто візуально важко відповісти на питання, чи є взаємозв'язок між цими ознаками: різними категоріями покупців та видами молочної продукції, тому необхідно дати аналіз розподілу частот у таблиці за рядками та графами. Будемо виходити з такого становища. Якщо ознака, покладена в основу угруповання за рядками (вид молочної продукції), не залежить від ознаки, покладеної в основу угруповання за стовпцями (тип покупця), то у кожному рядку (стовпці) розподіл частот має бути пропорційно розподілу їх у підсумковому рядку (стовпці). Такий розподіл можна розглядати як теоретичний (очікуваний), частоти якого розраховані у припущенні відсутності зв'язку між досліджуваними ознаками (сукупностями).

Розрахуємо очікувані частоти усередині таблиці 2.1.8 пропорційно розподілу частот у підсумковому рядку. Так, молоко як один з видів молочної продукції у залежності від поведінки відвідувачів ринку з частоти потрапляння у категорію покупців «практичних» та «імпульсивних» має наступні показники:

$$f_{11} = \frac{53 \times 80}{153} = 27,7; \quad (2.1.12)$$

$$f_{12} = \frac{53 \times 73}{153} = 25,3; \quad (2.1.12)$$

Для другого рядка, тобто для категорії сметана, ці показники мають вже такі значення:

$$f_{21} = \frac{55 \times 80}{153} = 28,8; \quad (2.1.12)$$

$$f_{22} = \frac{55 \times 73}{153} = 26,2; \quad (2.1.12)$$

Для третього рядка – категорія сир – наступні:

$$f_{31} = \frac{45 \times 80}{153} = 23,5; \quad (2.1.12)$$

$$f_{32} = \frac{45 \times 73}{153} = 21,5. \quad (2.1.12)$$

Отримані результати (обчислені значення частот) помістимо у таблицю 2.1.9.

Таблиця 2.1.9

**Дані щодо перспектив молочного бізнесу
з урахуванням очікуваних частот**

Вид молочної продукції	Очікувана частота переваг покупців:		Усього
	практичних	імпульсивних	
Молоко	27,7	25,3	53
Сметана	28,8	26,2	55
Сир	23,5	21,5	45
Усього	80	73	153

Розрахункове значення критерію χ^2 -квадрат визначимо за формулою:

$$\chi^2 = \sum_{i=1}^{k_1} \sum_{j=1}^{k_2} \frac{(f_{ij} - f_{ij}^*)^2}{f_{ij}^*}; \quad (2.1.15)$$

де: f_{ij} та f_{ij}^* – відповідно фактичні та теоретичні (очікувані) частоти у i -му рядку та у j -му стовпці;

k_1 та k_2 – відповідно число категорій у рядках та стовпцях таблиць 2.1.8 та 2.1.9.

Виконаємо відповідні розрахунки:

$$\begin{aligned} \chi_{\text{розрах}}^2 = & \frac{(38 - 27,7)^2}{27,7} + \frac{(15 - 25,3)^2}{25,3} + \frac{(24 - 28,8)^2}{28,8} + \frac{(31 - 26,2)^2}{26,2} + \frac{(18 - 23,5)^2}{23,5} \\ & + \frac{(27 - 21,5)^2}{21,5} = 12,4 \quad (2.1.15) \end{aligned}$$

Далі потрібно було порівняти розрахункове значення $\chi_{\text{розрах}}^2$ з табличним показником (зазвичай для рівня значущості 0,05 або 0,01). У розглянутому прикладі число ступенів свободи дорівнює двом, тобто $f = (3 - 1)(2 - 1) = 2$. Коли $\alpha = 0,05$ табличне значення $\chi_{\text{табл}}^2$ щодо $f = 2$ складає 5,991, коли $\alpha = 0,01$ відповідно 9,210. Оскільки $\chi_{\text{розрах}}^2 > \chi_{\text{табл}}^2$, то з упевненістю на 95 % (навіть на 99 %) можна стверджувати, що вплив психологічного типу покупця очевидним чином позначається на результатах комерційної діяльності кота Матроскіна.

Виявлення прихованих закономірностей
для вирішення бізнес-задач за допомогою Data-Mining

Проведення автоматичного аналізу даних передбачає використання Data-Mining (добування знань) – технологію інтелектуального аналізу даних щодо виявлення прихованих закономірностей у вигляді значущих особливостей кореляцій, тенденцій, шаблонів тощо.

Data-Mining – міждисциплінарна галузь, що виникла та розвинулась на основі статистики, теорії баз даних та штучного інтелекту. Технології Data-Mining гармонійно поєднує формалізовані та непараметричні методи аналізу, які впливають на її можливість розв'язати такі задачі: класифікацію (classification), кластеризацію (clustering), прогнозування (forecasting), асоціацію (associations), візуалізацію (visualization), аналіз та виявлення відхилень (analysis and identification of deviations), аналіз взаємозв'язків (correlation analysis), підведення підсумків (summarization) тощо.

Data-Mining складається з трьох стадій:

1. Виявлення закономірностей (вільний пошук).
2. Використання виявлених закономірностей для прогнозу невідомих значень (прогностичне моделювання).
3. Аналіз виключень (виявлення та пояснення аномалій, знайдених у закономірностях).

Data-Mining широко використовують щодо вирішення бізнес-задач у: банківській справі, сфері електронної комерції, страховій та фондовій діяльності, а також у маркетингових дослідженнях.

Контрольні питання

1. Що вивчає непараметрична статистика?
2. Назвіть переваги та недоліки непараметричних методів аналізу бізнес-процесів.
3. Як можна визначити наявність та напрям зв'язку за допомогою коефіцієнта кореляції рангів Спірмена?
4. Які коефіцієнти у бізнес-статистиці розраховують за допомогою «чотириклітинної» таблиці?
5. Що характеризує відношення шансів?

6. Статистичні показники аналізу якісних даних у бізнесі.
7. Що характеризує статистичний критерій χ^2 ?
8. Використання Data-Mining в аналізі бізнес-процесів.

2.2. . Статистичне оцінювання контролю якості та ефективності бізнес-процесів

Поняття якості та основні об'єкти її статистичної оцінки.

Загальна характеристика статистичних підходів до управління якістю та їх наочний аналіз.

Поняття ефективності та реінжинірингу бізнес-процесів.

Інформаційне забезпечення аналізу бізнес-процесів.

Методика статистичної оцінки малого бізнесу.

Поняття якості та основні об'єкти її статистичної оцінки

Статистичний контроль якості – це застосування статистичної методології щодо оцінювання якості продукції як складової загального менеджменту якості (TQM – Total Quality Management) та поліпшення результатів процесу виробництва у будь-якій сфері діяльності. Концепція TQM базується на тому, що постійне та одночасне повинно бути удосконалення трьох складових: 1) якості продукції, 2) якості організації процесів, 3) рівня кваліфікації персоналу. Це дає змогу швидше та ефективніше розвивати будь-який бізнес-процес.

Під бізнес-процесом розуміють будь-який вид економічної діяльності, який перетворює ресурси у продукти. Процес може складатися з інших процесів, званих підпроцесами, кожен з яких також є процесом. Статистичний контроль процесів полягає у використанні статистичних методів щодо контролю функціонування процесу таким чином, щоб можна було при необхідності регулювати процес та усувати неполадки або не втручатися у процес, якщо він працює належним чином. Усякий раз, коли є можливість з'ясувати, чому виникла та чи інша проблема, має місце не випадкова причина відхилень. Усі причини відхилень, які не заслуговують того, щоб їх визначати, об'єднують під назвою «випадкові причини відхилень». Якщо усі не випадкові причини відхилень визначені та

виключені, так що залишилися тільки випадкові причини, то кажуть, що процес знаходиться у стані статистичного контролю, або під контролем.

*Загальна характеристика статистичних підходів
до управління якістю та їх наочний аналіз*

Існують різні методи контролю якості продукції, серед яких особливе місце займають статистичні методи. Багато сучасних методів математичної статистики досить складні для сприйняття, а тим більше щодо широкого застосування всіма учасниками процесу управління якістю. Тому японські вчені відібрали з усієї безлічі групу методів, які найбільш застосовні у процесах контролю якості. Заслуга японців полягає у тому, що вони забезпечили простоту, наочність та візуалізацію цих методів, перетворивши їх в інструменти контролю якості, які можна зрозуміти та ефективно використовувати без спеціальної математичної підготовки. У той же час щодо усієї своєї простоти ці методи дозволяють зберегти зв'язок зі статистикою та дають можливість професіоналам у разі необхідності їх удосконалювати. Отже, до важливіших методів або інструментів контролю якості відносяться такі статистичні методи:

- контрольний листок,
- гістограма,
- діаграма розкиду,
- діаграма Парето,
- стратифікація (розшарування),
- діаграма Ісікави (причинно-наслідкова діаграма),
- контрольна карта.

Контрольний листок представляє собою паперовий бланк, на якому заздалегідь надруковані контрольовані параметри, згідно з якими можна заносити у листок дані за допомогою позначок або простих символів. Він дозволяє автоматично упорядкувати дані без їх подальшого переписування. Таким чином, контрольний листок – добрий засіб реєстрації даних. Кількість різних контрольних листків обчислюється сотнями, натомість щодо кожної конкретної мети може бути розроблено свій листок. Але алгоритм їх оформлення залишається незмінним.

Щодо складання контрольних листків, слід звернути увагу на те, щоб було зазначено, хто, на якому етапі процесу та протягом якого часу збирав дані, а також, щоб форма листка була простою та зрозумілою без додаткових пояснень. Важливо і те, щоб усі дані сумлінно фіксувалися та зібрана у контрольному листку інформація могла бути використана щодо аналізу бізнес-процесу.

Гістограма – це інструмент, що дозволяє візуально оцінити закон розподілу статистичних даних. Вона зазвичай будується щодо інтервальної зміни значення параметра. Тому на інтервалах, відкладених на осі абсцис, будують прямокутники (стовпчики), висоти яких пропорційні частотам інтервалів. На осі ординат відкладають абсолютні значення частот. Аналогічну форму гістограми можна отримати, якщо на осі ординат відкласти відповідні значення відносних частот. У цьому випадку сума площ усіх стовпчиків буде дорівнювати одиниці, що доволі зручно. Гістограма також дуже зручна для візуальної оцінки розташування статистичних даних у межах допуску. Щоб оцінити адекватність процесу вимогам споживача, ми повинні порівняти якість процесу з полем допуску, встановленим користувачем. Якщо є допуск, то на гістограму наносять верхню (SU) та нижню (SL) його межі у вигляді ліній, перпендикулярних осі абсцис, щоб порівняти розподіл параметра якості процесу з цими межами. Тоді можна побачити, чи добре розташовується гістограма у середині цих кордонів.

Діаграма розкиду – це інструмент, що дозволяє визначити вид та тісноту зв'язку між парами відповідних змінних.

Ці дві змінні можуть ставитися щодо:

- характеристики якості та впливу на неї факторів,
- двох різних характеристик якості,
- двох факторів, що впливають на одну характеристику якості.

Щодо виявлення зв'язку між ними і служить діаграма розкиду, яку також ще називають полем кореляції. Використання діаграми розкиду у процесі контролю якості не обмежується тільки виявленням виду та тісноти зв'язку між парами змінних. Діаграма розкиду використовується також щодо виявлення причинно-наслідкових зв'язків показників якості та факторів, що на них впливають.

Діаграма Парето показує причини різних недоліків, які розташовані у порядку від найбільш часто виникаючих до найменш часто виникаючих, що дозволяє зосередити увагу на найважливіших проблемах. На діаграмі Парето найвищий стовпчик буде завжди зліва (показуючи найбільш часто виникаючі проблеми), а найнижчий – справа. Лінія кумулятивного відсотка завжди спрямована вгору і щодо руху вправо має тенденцію наближатися до горизонтальної лінії.

Розрізняють два види діаграм Парето:

1. Діаграма Парето за результатами діяльності, яка призначена щодо виявлення головної проблеми та відображає наступні небажані результати діяльності:

- якість – дефекти, поломки, помилки, відмови, рекламации, ремонти, повернення продукції;
- собівартість – обсяг втрат та витрати;
- терміни поставок – нестача запасів, помилки у складанні рахунків, зрив термінів постачань;
- безпека – нещасні випадки, трагічні помилки та аварії.

2. Діаграма Парето з наступних підстав, яка відображає причини проблем, що виникають у ході виробництва і використовується для виявлення головної з них:

- виконавець роботи – зміна, бригада, вік, досвід роботи, кваліфікація, індивідуальні характеристики;
- обладнання – верстати, агрегати, інструменти, оснащення, організація використання, моделі, штампи;
- сировина – виробник, вид сировини, завод-постачальник, партія;
- метод роботи – умови виробництва, замовлення-наряди, прийоми роботи, послідовність операцій;
- вимірювання – точність (вказівок, читання, приладова), вірність та повторюваність (вміння дати однакову вказівку у наступних вимірах одного і того ж значення), стабільність (повторюваність протягом тривалого періоду), спільна точність, тобто разом з приладовою точністю та таруванням приладу, тип вимірювального приладу (аналоговий або цифровий).

Одним з найбільш ефективних статистичних методів, які широко використовуються у системі управління якістю, є метод *стратифікації або розшаровування*. Відповідно до цього методу вводять розшарування статистичних даних, тобто групують дані у залежності від умов їх отримання та роблять обробку кожної групи даних окремо. Дані, розділені щодо груп відповідно до їх особливостей, називають шарами (стратами), а сам процес поділу на верстви (страти) – розшаруванням (стратифікацією).

Метод розшаровування досліджуваних статистичних даних – це інструмент, що дозволяє зробити селекцію даних, яка відображатиме необхідну інформацію про процес.

Існують різні методи розшаровування, застосування яких залежить від конкретних завдань. Наприклад, дані, що відносяться до виробу, який зроблено у цеху на робочому місці, можуть у якійсь мірі відрізнятися у залежності від виконавця, використовуваного обладнання, методів проведення робочих операцій, температурних умов та т. п. Усі ці відмінності можуть бути факторами розшаровування. У виробничих процесах часто використовується метод *5М*, що враховує фактори, які залежать від людини (man), машини (machine), матеріалу (material), методу (method), вимірювання (measurement).

За якими ж критеріями можна виконувати розшарування? Воно може здійснюватися за наступними критеріями:

- за виконавцями – за кваліфікацією, стажу роботи та т. п.;
- за машинами й устаткуванням – за новим чи старим обладнанням, маркою, конструкцією, яка випускає фірма, і т. п.;
- за матеріалом – за місцем виробництва, фірмою-виробником, партією, якістю сировини і т. п.;
- за способом виробництва – з температури, технологічного прийому, місця виробництва і т. п.;
- за вимірюванням – за методом вимірювання, типом вимірювальних засобів або їх точності і т. п.

Однак користуватися цим методом не так просто. Іноді розшарування за, здавалося б, очевидним параметром не дає очікуваного результату.

У цьому випадку потрібно продовжити аналіз даних з інших можливих параметрів у пошуках вирішення виниклої проблеми.

Діаграма Ісікави. Результат процесу залежить від численних факторів, між якими існують відносини типу причина – наслідок (результат). Діаграма причин та наслідків – це засіб, що дозволяє висловити ці відносини у простій та доступній формі.

У 1953 році професор Токійського університету Каору Ісікава, обговорюючи проблему якості одного заводу, підсумовував думку інженерів у формі діаграми причин та результатів. Коли діаграму почали застосовувати щодо практики, вона була дуже корисною і скоро стала широко використовуватися у багатьох компаніях Японії, отримавши назву діаграми Ісікава. Вона була запропонована до включення у японський промисловий стандарт (JIS) з термінологією у галузі контролю якості та визначенням у ньому наступним чином: діаграма причин та наслідків, яка показує відношення між показником якості та впливаючим на нього фактором. Причинно-наслідкова діаграма – це інструмент, що дозволяє виявити найбільш істотні фактори (причини), що впливають на кінцевий результат (наслідок).

Якщо у результаті процесу якість виробу виявилася незадовільною, то у системі причин, тобто у якійсь точці процесу відбулося відхилення від заданих умов. Якщо ця причина може бути виявлена й усунена, то будуть випускатися вироби тільки високої якості. Більш того, якщо постійно підтримувати задані умови процесу, то можна забезпечити формування високої якості виробів, що випускаються. Важливо також, що отриманий результат – показники якості (точність розмірів, ступінь чистоти, значення електричних величин і т.п.) – виражається конкретними даними. Використовуючи ці дані, за допомогою статистичних методів здійснюють контроль процесу, тобто перевіряють систему причинних факторів. Таким чином, процес контролюється за фактором якості.

Як зібрати дані, необхідні для побудови діаграми Ісікава? Інформація щодо показників якості для побудови діаграми збирається з усіх доступних джерел з використанням: журналу реєстрації операцій, журналу реєстрації даних поточного контролю, повідомлення робочих

виробничої дільниці і т. п. При побудові діаграми вибираються найбільш важливі з технічної точки зору чинники. Щодо цієї мети широко використовується експертна оцінка. Дуже важливо простежити кореляційну залежність між причинними факторами (параметрами процесу) та показниками якості. У цьому випадку параметри легко піддаються кореляції. Щодо цього, аналіз дефектів виробів слід розділити на випадкові та систематичні, звернувши особливу увагу на можливість виявлення та подальшого усунення у першу чергу причини систематичних дефектів.

Важливо пам'ятати, що показники якості, які є наслідком процесу, обов'язково відчують розкид. Пошук чинників, що мають особливо великий вплив на розкид показників якості виробу (тобто на результат), називають дослідженням причин.

Карта контролю містить послідовність вимірів параметрів процесу, центральну лінію та контрольні кордону, які обчислюють, щоб приймати рішення про те, чи знаходиться процес під контролем чи ні? Якщо приймається рішення про те, що процес не знаходиться під контролем, карта контролю допомагає з'ясувати та усунути відповідну проблему. Перевіряються гіпотези, що формуються наступним чином:

H_0 – процес знаходиться під контролем;

H_1 – процес не знаходиться під контролем.

Рівень помилкової тривоги характеризує, як часто втручання у процес відбувається без необхідності. Це поняття аналогічно поняттю помилки першого роду у перевірці гіпотез. Загальноприйняті карти контролю будуються, виходячи не з двох стандартних помилок параметрів, оскільки звичайний для статистичної перевірки гіпотез рівень 5 % помилки першого роду неприпустимо високий щодо більшості додатків контролю якості. Коли процес знаходиться під контролем, графік карти контролю не виходить за межі. Якщо будь-який вимір виходить за межі контрольних меж, або вище верхньої межі, або нижче нижньої межі, то роблять висновок, що процес вийшов з-під контролю. Раптовий стрибок значень у графіку карти контролю на інший рівень або поступовий зсув точок вгору або вниз може також свідчити про те, що процес не

знаходиться під контролем, навіть незважаючи на те, що усі точки розташовані у межах контрольних меж.

Для складання карти кількісних вимірів зазвичай вибирають невеликий розмір вибірки ($n = 4$ або $n = 5$) і потім наносять на карту результати багатьох послідовних вибірок. *X*-карта містить середнє значення кожної вибірки, відповідну центральну лінію і контрольні кордону, що дозволяє контролювати рівень процесу. *R*-карта містить розмах (найбільше значення мінус найменше) значень для кожної вибірки, центральну лінію і контрольні кордону, що дозволяє контролювати мінливість процесу.

Відсоткова карта показує відсоток бракованих виробів, відповідну центральну лінію і контрольні кордону, що дозволяє контролювати рівень, на якому процес починає виробляти неякісні вироби. Для побудови такої карти розмір вибірки повинен бути більше, ніж для карти кількісних виробів. Емпіричне правило полягає у тому, що очікувана кількість бракованих виробів у вибірці має бути не менше п'яти.

Поняття ефективності та реінжинірингу бізнес-процесів

Ефективність в економіці часто асоціюють з оптимальністю цілеспрямованої дії, способу, механізму реалізації або стану суб'єкта за наявності альтернатив. Вона проявляється через співвідношення між метою та отриманими результатами, використаними ресурсами, цілями, обставинами зовнішнього середовища та часовими рамками. Що ж стосується ефективності бізнес-процесів, то вона показує, наскільки добре виконуються процеси, які тільки через удосконалення можуть бути більш ефективними.

Покращення бізнес-процесів і відповідно їх ефективності, здійснюється шляхом *реінжинірингу*. Родоначальником терміну вважають Майкла Хаммера [4, с. 245]. За його визначенням, реінжиніринг – це фундаментальне переосмислення і радикальне перепроєктування бізнес-процесів з метою досягнення істотного поліпшення ефективності їх функціонування. Досягти оцінки функціонування процесу можливо лише на основі досконального бізнес-аналізу.

Бізнес-аналіз повинен бути комплексним дослідженням внутрішніх і зовнішніх чинників, що впливають на процес функціонування бізнес-процесів підприємства, оскільки основна мета проведення аналізу – підвищення ефективності функціонування бізнес-процесів та прийняття оптимальних управлінських рішень по забезпеченню їх стійкого розвитку.

Для успішної реалізації мети аналізу бізнес-процесів необхідно вирішити наступні завдання:

- визначення економічної ефективності використання ресурсів бізнес-процесів;
- оцінка виконання планів, прогнозів, управлінських рішень за ефективним використанням економічного потенціалу бізнес-процесів підприємства;
- вивчення впливу об'єктивних і суб'єктивних, зовнішніх і внутрішніх чинників на результати бізнес-процесів підприємства;
- розробка й обґрунтування заходів, спрямованих на оптимізацію бізнес-процесів підприємства і прийняття ефективних управлінських рішень.

Інформаційне забезпечення аналізу бізнес-процесів

Необхідною умовою якісного проведення аналізу бізнес-процесів є його інформаційне забезпечення. Інформаційне забезпечення аналізу – це сукупність інформаційних ресурсів, необхідних для досягнення цілей аналізу. Як інформаційні ресурси виступає інформація, що циркулює в системі управління і є основою управлінських рішень.

Інформаційне забезпечення бізнес-аналізу можна розділити на наступні основні блоки:

1. Нормативно-довідкова – законодавчі акти, нормативи урядових і старанних структур, норми, нормативи, ліміти та інші регламенти бізнес-процесів.
2. Планова інформація – міститься у бізнес-плані підприємства або у плані економічного (соціального) розвитку підприємства.
3. Фактографічна інформація – сукупність даних економічного характеру про факти, які дійсно здійснювалися у бізнес-процесах

підприємства. Ця інформація, як правило, формується обліково-фінансовою службою. Фактографічну інформацію можна підрозділити на три блоки:

- облікова інформація, що міститься в облікових документах підприємства – первинних документах, регістрах бухгалтерського обліку, відомостях та інше;
- звітна інформація, що міститься у статистичній та фінансовій звітності, наприклад, форма 1 «Баланс», форма 2 «Звіт про фінансові результати», «Звіт про працю» і т. п. Саме цей вид інформації має основне значення при проведенні зовнішнього аналізу;
- позаоблікова інформація, що міститься в актах перевірок, аудиторських висновках, а також у документах, регулюючих бізнес-процеси підприємства (офіційні документи, договори, угоди, рекламації, технічна і технологічна інформація);
- нормативно-довідкова інформація, яка регулює усі бізнес-процеси підприємства, це: Закон України «Про бухгалтерський облік та фінансову звітність»; Положення (стандарти) бухгалтерського обліку в Україні; господарсько-правові документи; технічна й технологічна документація; інформація, отримана в ході маркетингових досліджень тощо.

Джерелами інформації для аналізу процесу постачання матеріальних ресурсів та процесу розрахунків з покупцями виступають:

- довідкова інформація: плановані обсяги запасів та обсяги надходження (портфель замовлень), план матеріально-технічного постачання, нормативи товарних запасів, нормативи товарних витрат;
- облікова інформація: це накладні на відвантаження продукції, товаротransпортні накладні, рахунки, журнал реєстрації виставлених рахунків-фактур, оперативні дані відділу матеріально-технічного постачання, відомості, звіти матеріально-відповідальних осіб, реєстри синтетичного обліку за рахунками «Каса», «Товари», «Розрахунки з постачальниками»;

- звітна інформація: форми № 1 «Баланс», форма № 2 «Звіт про фінансові результати», форми статистичної звітності № 1-СН, 3-СН, 4-СН, 11-СН, 12-СН про наявність і використання матеріальних ресурсів, форма № 5-з про витрати на виробництво;
- позаоблікова інформація: акти експертів, протоколи зборів, договори, угоди, контракти на постачання сировини та матеріалів.

Джерелами інформації для аналізу процесу реалізації готової продукції є:

- довідкова інформація – це індекси цін на споживчі товари, нормативи товарних запасів, нормативи товарних витрат, нормативи обороту на торговій площі;
- планова інформація – це дані бізнес-плану, плану реалізації у розрізі товарних груп, плану постачання товарів;
- облікова інформація – це накладні на відвантаження продукції, товаротранспортні накладні, рахунки-фактури, журнал реєстрації виставлених рахунків-фактур, книга продажів, відомості, звіти матеріально-відповідальних осіб, реєстри синтетичного обліку по рахункам «Каса», «Товари», «Розрахунки з постачальниками»;
- звітна інформація – це форми № 1 «Баланс», форма № 2 «Звіт про фінансові результати», форма № 3 – торг (коротка) місячна, форма № 7 – торг (роздріб), «Звіт про наявність торгової мережі» – річна, форма № 1 – торг «Звіт про товарообіг»;
- позаоблікова інформація – це акти експертів, протоколи зборів, договір на поставку товарів.

У зв'язку з особливостями етапу становлення малого бізнесу країни і недостатнім рівнем розвитку статистики малого бізнесу у деяких випадках використовуються показники, які можуть дати опосередковану інформацію:

1. Показники, що відображають стан сектора малого та середнього бізнесу (МСБ) у цілому та за окремими галузями:

- розподіл числа малих та середніх підприємств за галузями економіки;
- розподіл показника «частка малих / середніх підприємств у загальній кількості підприємств в регіоні» за галузями економіки;

- розподіл показника «кількість малих / середніх підприємств на 1000 чоловік населення регіону» за галузями економіки;
- розподіл середньої чисельності працівників малих та середніх підприємств за галузями економіки;
- розподіл показника «середня чисельність працівників на одному малому / середньому підприємстві» за галузями економіки;
- розподіл обсягу виробництва продукції (робіт, послуг) малих підприємств за галузями економіки.

2. Показники, що відображають ефективність функціонування сектора МСБ в цілому та за окремими галузями:

- розподіл виручки від реалізації продукції (робіт, послуг) малих та середніх підприємств, зокрема у розрахунку на одне підприємство, за галузями економіки;
- розподіл прибутку малих та середніх підприємств, зокрема у розрахунку на одне підприємство, за галузями економіки;
- розподіл показника «частка прибуткових малих / середніх підприємств у галузі» за галузями економіки.

3. Показники ефективності використання трудових ресурсів сектора МСБ у цілому та за окремими галузями:

- розподіл продуктивності праці малих і середніх підприємств (обсяг виробництва у розрахунку на одного працівника підприємства) за галузями економіки;
- розподіл прибутку малих і середніх підприємств у розрахунку на одного працівника підприємства за галузями економіки.

4. Показники динаміки малих підприємств у цілому і за окремими галузями:

- розподіл темпів зростання числа малих підприємств, чисельності працівників і обсягу виробництва продукції (робіт, послуг) малих підприємств за галузями економіки.

Статистичні методи передбачають широке застосування абсолютних і відносних показників. Для дослідження необхідно використовувати декілька типів відносних показників з метою вивчення особливостей функціонування і розвитку малого бізнесу:

- показники інтенсивності розвитку характеризують поширення

суб'єктів малого бізнесу у просторі (забезпеченість малих підприємств у розрахунку на 10 тис. осіб населення);

- показники співвідношень засвідчують про співвідношення одних показників до інших.

Методика статистичної оцінки малого бізнесу

Як зазначалося раніше, для якісного оцінювання ефективності бізнес-процесів необхідно розробити систему показників, які б розкривали інформацію про досліджувані бізнес-процеси відповідно до специфіки діяльності підприємства. У нашому прикладі розглянемо методику оцінювання підприємств малого бізнесу. Вона містить:

Визначення змін у малому бізнесі. Для цього використовуються показники:

- абсолютний приріст (спад);
- темпи зростання (спаду);
- коефіцієнт дивергенції:

$$K_d = \frac{\Delta B_{p-n}}{\Delta B_{країна}}; (2.2.1)$$

де: ΔB_{p-n} – приріст виробництва продукції (робіт, послуг) малих підприємств на одного жителя регіону за певний період;

$\Delta B_{країна}$ – приріст виробництва продукції (робіт, послуг) малих підприємств на одного жителя країни за цей період.

Інтегральна оцінка розвитку малих підприємств базується на розрахунку індексу рівня розвитку малих підприємств, який враховує добуток часткових індексів:

$$I_{pМП} = I_{кМП} \times I_{зМП} \times I_{нМП}; (2.2.2),$$

де: $I_{pМП}$ – інтегральний індекс розвитку малих підприємств;

$I_{кМП}$ – індекс кількості малих підприємств (частка малих підприємств у загальній кількості підприємств);

$I_{зМП}$ – індекс зайнятості на малих підприємствах (частка зайнятих на МП до загальної кількості зайнятих);

$I_{нМП}$ – індекс обсягів виробленої продукції малими підприємствами (частка малих підприємств у загальному обсязі виробленої продукції (робіт, послуг) в країні).

Визначення реальних масштабів діяльності малих підприємств країни. Враховується тіньовий компонент діяльності за умов зіставлення фактичних основних показників діяльності малих підприємств («реальних» або «потенційних»). «Реальний» дохід від реалізації продукції (робіт, послуг) D_p обчислюють за формулою:

$$D_{pi} = K_{pi} \times D_i; (2.2.3),$$

де: K_{pi} – коефіцієнт «реальної» діяльності і-го року;

D_i – фактичний дохід від реалізації продукції (робіт, послуг) і-го року.

Коефіцієнт «реальної діяльності» обчислюють як співвідношення кількості зареєстрованих суб'єктів, зокрема малі підприємства, до кількості діючих:

$$K_{pi} = \frac{Z_{mpi}}{D_{mpi}}; (2.2.4),$$

де: Z_{mpi} – кількість зареєстрованих малих підприємств і-го року;

D_{mpi} – кількість діючих малих підприємств і-го року (які звітували).

Для оцінки впливу чинників становлення і розвитку малого бізнесу використовують метод побудови множинної регресії. Загальне призначення моделі множинної регресії полягає в аналізі зв'язку між декількома незалежними змінними (пояснювальними факторами) та залежною змінною (результативним фактором). Для типізації одиниць адміністративно-територіального поділу країни за рівнем та характером розвитку малого бізнесу використовуються методи факторного та таксономічного аналізу.

Потенційні кредитори та інвестори для гарантування прибутку на вкладений капітал детально аналізують фінансово-господарську діяльність позичальника. Одним із важливих етапів є дослідження фінансової стійкості та стабільності підприємства-позичальника, що включає такі показники:

коефіцієнт фінансової незалежності ($K_{ф.н.}$);

коефіцієнт фінансової стабільності ($K_{ф.с.}$);

коефіцієнт фінансової залежності ($K_{ф.з.}$);

коефіцієнт фінансового ризику ($K_{ф.р.}$);

коефіцієнт забезпечення власними коштами ($K_{з.к.}$).

Ринкову стійкість підприємства характеризує коефіцієнт фінансової незалежності ($K_{ф.н.}$):

$$K_{ф.н.} = \frac{\text{Власний капітал}}{\text{Вартість майна підприємства}}; (2.2.5)$$

Коефіцієнт незалежності бажано підтримувати на достатньо високому рівні, що дасть можливість забезпечити стабільну структуру капіталу. Саме такій структурі віддають перевагу інвестори та кредитори.

На практиці питання оптимального співвідношення власного і позикового капіталу не вирішується однозначно. Кожне підприємство визначає його у взаємозв'язку зі структурою основного й оборотного капіталу, швидкістю оборотності обігових коштів та інших показників.

Коефіцієнт фінансової стабільності ($K_{ф.с.}$):

$$K_{ф.с.} = \frac{\text{Власний капітал}}{\text{Позиковий капітал}}; (2.2.6)$$

Значення цього коефіцієнта не повинно перевищувати 1.

Коефіцієнт фінансової залежності ($K_{ф.з.}$):

$$K_{ф.з.} = \frac{\text{Позиковий капітал}}{\text{Загальний капітал}}; (2.2.7)$$

Він характеризує залежність діяльності підприємства від позикового капіталу. Оптимальним значенням цього коефіцієнта вважається 0,4.

Коефіцієнт фінансового ризику ($K_{ф.р.}$):

$$K_{ф.р.} = \frac{\text{Позиковий капітал}}{\text{Власний капітал}}; (2.2.8)$$

Це – найвагоміший показник, який свідчить про фінансову незалежність (автономність) підприємства від залучення позикових коштів. Він показує, скільки позикових коштів залучає підприємство на 1 гривню власного капіталу.

Оптимальне значення $K_{ф.р.}$ – 0,3–0,5. Критичним значенням цього коефіцієнта вважається одиниця. Однак, економісти-статистики вважають, що за високих показників оборотності обігових коштів критичне значення коефіцієнта може перевищувати одиницю без істотних наслідків для ринкової стійкості.

Малі та середні підприємства майже усіх галузей є досить залежними від позикових коштів. Для вирішення питання щодо забезпечення підприємства власними коштами розраховується коефіцієнт забезпечення власними коштами ($K_{з.к.}$):

$$K_{з.к.} = \frac{\text{Власний капітал} - \text{Необоротні активи}}{\text{Оборотний капітал}}; (2.2.9)$$

Значення цього коефіцієнта повинно бути більше, ніж 0,1.

Висновок: Формування ринкової системи господарювання в Україні пов'язане із зростанням підприємницької активності в усіх сферах економіки. Один із перспективних напрямів створення конкурентно-ринкового середовища – розвиток малого та середнього бізнесу. Світовий досвід і практика господарювання показують, що найважливішим компонентом ринкової економіки є існування, взаємодія та оптимальне співвідношення великих, середніх та малих підприємств, а статистична оцінка ефективності бізнес-процесів у середніх та малих підприємствах є дуже важливою для з'ясування того, чи вірно функціонують ці підприємства.

Контрольні питання

1. Які основні завдання статистичного вивчення якості?
2. Назвіть сім основних методів або інструментів контролю якості.
3. З якою метою будують діаграму Парето? Схарактеризуйте алгоритм її побудови.
4. Що таке діаграма Ісікави? Як її будують і використовують у статистичному вивченні якості продукції.
5. Поясніть суть понять «ефективність» та «реінжиніринг бізнес-процесів».
6. Завдання статистичного оцінювання ефективності бізнес-процесів.
7. Алгоритм інформаційного забезпечення бізнес-аналізу, що циркулює у системі управління.
8. Методика оцінки ефективності малого бізнесу.

2.3. Практичний тренінг розделу 2

Тести щодо самоконтролю

Непараметричні методи у бізнес-статистиці

1. Непараметрична статистика – це:
А) самостійна галузь математичної статистики;
В) класичний метод аналізу даних;
С) статистична перевірка гіпотез;
Д) усі відповіді вірні.
2. Статистичний метод, який доволі простий та якому притаманна стійкість висновків:
А) коефіцієнт кореляції рангів Спірмена;
В) квадратична спряженість Пірсона χ^2 ;
С) коефіцієнт спряженості Чупрова;
Д) усі відповіді невірні.
3. Ранги оцінки восьми підприємств іноземними інвесторами та зовнішніми незалежними експертами у сфері інвестування розподілені так:

Наданий ранг	
Іноземними інвесторами	Зовнішніми незалежними експертами
8	8
7	4
6	1
4	7
1	3
2	6
3	2
5	5

Коефіцієнт рангової кореляції Спірмена дорівнює:

А) 0,24; В) 0,87; С) 0,35; Д) 0,51.

4. Коефіцієнти асоціації та контингенції використовують для оцінки зв'язку у бізнес-статистиці між:
А) атрибутивними ознаками;
В) кількісними ознаками;
С) альтернативними ознаками;
Д) ознаками порядкової шкали.

5. Розподіл молодих працівників за ступенем задоволеності умовами праці та професійною мобільністю характеризується такими даними:

Ступінь задоволеності умовами праці	Мають намір змінити професію	Не мають наміру змінювати професію	Разом
Задоволені	10	30	40
Незадоволені	20	5	25
Разом	30	35	65

Коефіцієнт контингенції становить:

- А) $-0,54$; В) $-0,84$; В) $0,25$; Д) $0,47$.

6. Коефіцієнт спряженості Чупрова розраховується за формулою:

А) $K = \frac{\sum_1^n xy - n\bar{x}\bar{y}}{n\sqrt{\sigma_x^2 \sigma_y^2}}$; В) $K = \sqrt{\frac{\chi^2}{n\sqrt{(k_1-1) \cdot (k_2-1)}}}$; С) $K = \frac{\sum C - \sum H}{\sum C + \sum H}$; Д) $K = \frac{2S}{n(n-1)}$.

7. Коефіцієнт спряженості Крамера розраховується за формулою:

А) $K = \frac{2S}{n(n-1)}$; В) $K = \frac{\sum C - \sum H}{\sum C + \sum H}$; С) $K = \frac{\sum_1^n xy - n\bar{x}\bar{y}}{n\sqrt{\sigma_x^2 \sigma_y^2}}$; Д) $K = \sqrt{\frac{\chi^2}{n(m_{\min}-1)}}$.

Аналіз закономірностей для якісних даних

8. Для обробки масивів якісних даних часто використовують показники:

- А) статистичні;
В) комбіновані;
С) евристичні;
Д) усі відповіді вірні.

9. Щодо аналітичного оцінювання якісних даних використовують наступні показники:

- А) моду та середній нормований бал;
В) дисперсію альтернативної ознаки;
С) коефіцієнти локалізації та концентрації;
Д) коефіцієнти взаємного спряження та рангової кореляції.

10. Показник, який вимірює різницю між частотами, що спостерігаються (експериментальними) f_e , та очікуваними (теоретичними) частотами f_m , називається:

- А) « χ^2 -квадрат»;
- В) коефіцієнт Пірсона;
- С) коефіцієнт Чупрова;
- Д) коефіцієнт детермінації.

11. Значення коефіцієнтів Чупрова, Крамера та Пірсона у бізнес-статистиці змінюється у межах від:

- А) $[0, 1]$; В) $[-1, 1]$; С) $[-\infty, +\infty]$; Д) $[0, +\infty]$

12. Якщо виявляється справедливим нерівність $\chi^2_{\text{розр}} > \chi^2_{\text{крит}}$, то із заданою вірогідністю (або рівнем значущості) можна стверджувати, що частоти, котрі спостерігаються:

- А) незначно відрізняються від опорних показників, отже, наші фактичні результати не мають значущих відмінностей від заданих опорних значень;
- В) значимо відрізняються від тих, які очікуються, виходячи з відомих нам опорних значень відсотків (частот);
- С) зовсім не відрізняються;
- Д) усі відповіді вірні.

13. При обчисленні статистики « χ^2 -квадрат» для перевірки гіпотези про незалежність, число ступенів свободи розраховують за такою формулою:

- А) число ступенів свободи = (Число категорій для першої змінної -2) * (Число категорій для другої змінної -2);
- В) число ступенів свободи = (Число категорій для першої змінної -1) + (Число категорій для другої змінної -1);
- С) число ступенів свободи = (Число категорій для першої змінної -1) / (Кількість категорій для другої змінної -1);
- Д) число ступенів свободи = (Число категорій для першої змінної -1) * (Число категорій для другої змінної -1).

14. Data-Mining – технологія інтелектуального аналізу даних щодо виявлення прихованих закономірностей у вигляді значущих особливостей кореляцій, тенденцій, шаблонів тощо, виникла та розвинулась на основі:

- А) статистики;
- В) теорії баз даних;
- С) штучного інтелекту;
- Д) усі відповіді вірні.

Статистичне оцінювання контролю якості

- 15.** Теоретичні та прикладні проблеми оцінювання якості об'єктів (виробів, послуг, процесів, систем) вивчає наука:
- А) кваліметрія;
 - В) економетрія;
 - С) темпологія;
 - Д) носеологія.
- 16.** Концепція TQM (Total Quality Management) базується на необхідності постійного та одночасного удосконалення трьох складових:
- А) якості продукції;
 - В) якості організації процесів;
 - С) рівня кваліфікації персоналу;
 - Д) якості ресурсів.
- 17.** Відповідно до принципів менеджменту якості об'єтами вимірювання мають виступати:
- А) задоволення споживача;
 - В) процеси діяльності;
 - С) продукти як результат діяльності;
 - Д) ресурси діяльності.
- 18.** Визначити оптимальні значення параметрів якості продукції та процесів її створення дає змогу:
- А) експертний контроль;
 - В) статистичний приймальний контроль;
 - С) статистичне управління процесами;
 - Д) планування експерименту.
- 19.** Відповідність створеного продукту встановленим стандартам чи вимогам споживача дає змогу:
- А) діаграма Парето;
 - В) контрольна карта;
 - С) стратифікація (розшарування);
 - Д) діаграма Ісікави (причинно-наслідкова діаграма).
- 20.** За відношенням до властивостей продукції розрізняють такі групи показників якості:

- А) надійності та естетичності;
- В) ергономічності та органолептичності;
- С) економічності;
- Д) інструментальності та експертності.

21. Ергономічні показники якості продукції характеризують:

- А) здатність виробу до передавання краси у наочній формі;
- В) конструктивно-технологічні рішення у виробництві продукції;
- С) відповідність виробу фізіологічним та психологічним властивостям споживача;
- Д) відповідність виробу гігієнічним умовам життєдіяльності та працездатності людини при взаємодії її з виробом.

Статистичне оцінювання ефективності бізнес-процесів

22. Покращення бізнес-процесів, тобто підвищення їх ефективності, здійснюється шляхом їх:

- А) модернізації;
- В) автоматизації;
- С) реінжинірингу;
- Д) реструктуризації.

23. Родоначальником терміну «реінжиніринг» вважається:

- А) К. Чупров;
- В) Л. Козерод;
- С) Д. Антіпов;
- Д) М. Хаммер.

24. Щодо успішної реалізації мети аналізу бізнес-процесів необхідно:

- А) визначення економічної ефективності використання ресурсів бізнес-процесів;
- В) оцінка виконання планів, прогнозів, управлінських рішень тощо;
- С) вивчення впливу об'єктивних і суб'єктивних, зовнішніх і внутрішніх чинників на результати бізнес-процесів підприємства;
- Д) розробка й обґрунтування заходів, спрямованих на оптимізацію бізнес-процесів підприємства і прийняття ефективних управлінських рішень.

- 25.** Інформаційне забезпечення бізнес-аналізу можна розділити на:
- А) нормативно-довідкове;
 - В) планово-інформаційне;
 - С) фактографічно-інформаційне;
 - Д) звітно-облікове тощо.
- 26.** У якості основних складових системи показників, які характеризують бізнес-процеси підприємства, виділяють наступні групи:
- А) показники задоволеності споживачів результатами діяльності підприємства;
 - В) показники результативності бізнес-процесів;
 - С) показники вартості бізнес-процесів;
 - Д) показники фрагментації.
- 27.** Методика оцінки малого бізнесу містить такі складові:
- А) визначення змін у малому бізнесі;
 - В) інтегральна оцінка розвитку малих підприємств;
 - С) визначення реальних масштабів діяльності малих підприємств країни;
 - Д) усі відповіді вірні.
- 28.** В сучасній фаховій літературі, залежно від масштабів проведення реінжинірингу, вчені виділяють такі його базові види:
- А) «разові поліпшення»;
 - В) «кризовий»;
 - С) «адаптивний»;
 - Д) «тотальне моделювання».

РЕЗЮМЕ

АНАЛІТИЧНІ МОЖЛИВОСТІ СТАТИСТИКИ

Аналітична статистика – це наука збору, аналізу, представлення й інтерпретації даних. Державні потреби в даних перепису, а також інформація про різні економічні заходи забезпечили більшу частину раннього стимулу в області статистики. В даний час необхідність перетворення великого обсягу даних, доступних у багатьох прикладних областях в корисну інформацію, стимулювала як теоретичний, так і практичний розвиток статистики. Особливість статистики полягає в тому, що ця наука несе нам інформацію у кількісному вигляді, тобто статистика розмовляє на мові чисел, тим самим відображає соціальне та економічне життя у всьому різноманітті його форм і проявів. Статистична інформація – це сукупність статистичних даних, що відображає соціально-економічні явища та процеси, які формуються в процесі статистичного спостереження, а потім систематизуються, групуються, аналізуються та узагальнюються.

Бізнес-аналітика – це процес аналізу інформації для прийняття бізнесових рішень. Вона включає методи збору і обробки інформації, оцінку ризиків, моделювання і прогнозування за допомогою інформаційних і телекомунікаційних технологій. Головне завдання бізнес-аналітики – пропонувати і вносити зміни, які принесуть користь організації. Бізнес-аналітика здатна на основі аналізу конкретних показників бізнес-процесів обчислити області, зміни в яких матимуть позитивний потенційний вплив на загальний результат. Бізнес-аналітика займається аналізом внутрішньокорпоративних процесів, вивчає роботу компанії, намагається мінімізувати витрати і підвищити ефективність і результативність організації. Важливим інструментом бізнес-аналітики є бізнес-аналіз, який у свою чергу поділяється на своєрідні розділи та етапи. Узагальнені розділи бізнес-аналізу схематично наведені на рис. 1, етапи проведення бізнес-аналізу наведені на рис. 2. Дані схеми корисно використовувати щодо бізнес-аналізу у будь-якій галузі економіки. Це наукові підходи до аналітичних можливостей саме статистики.

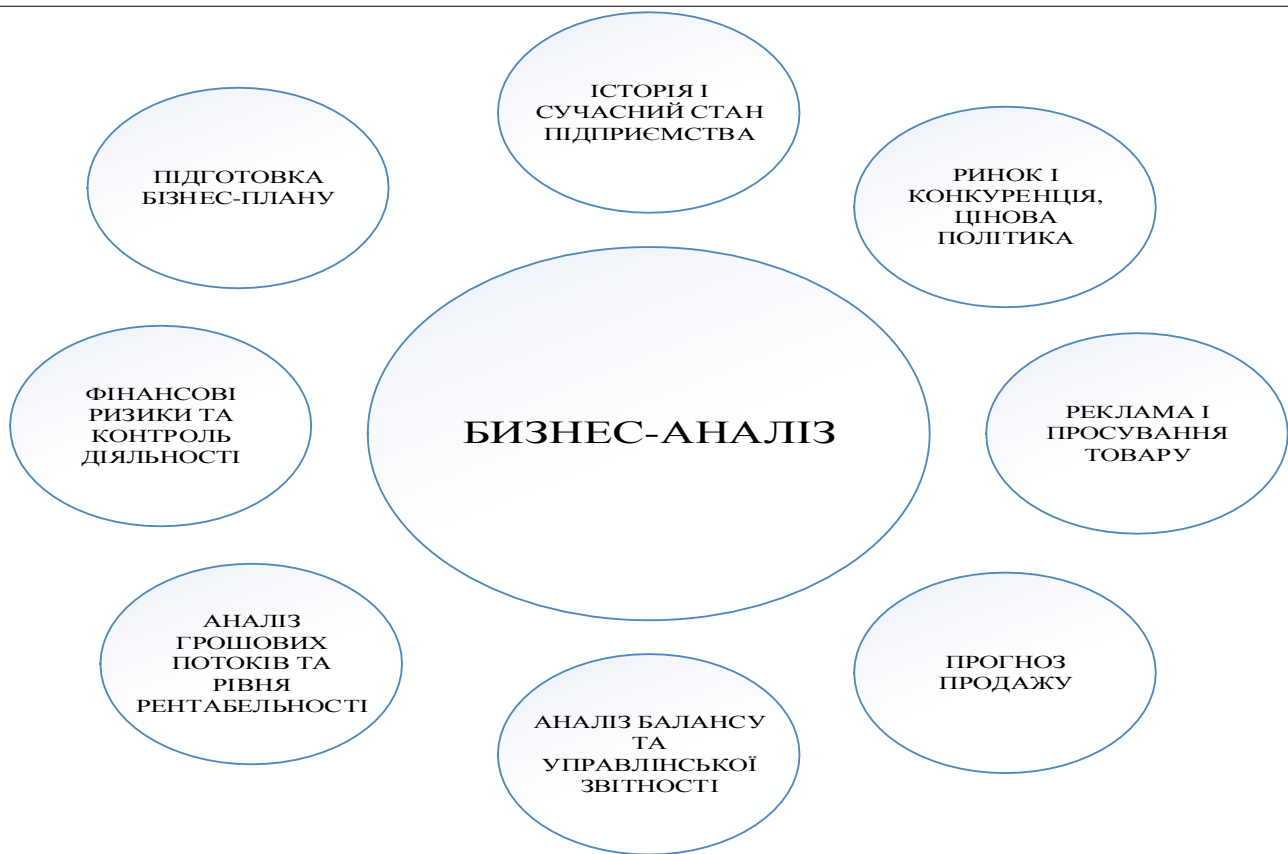


Рис. 1. Узагальнені розділи бізнес-аналізу

Джерело: Побудовано автором за матеріалами [30].

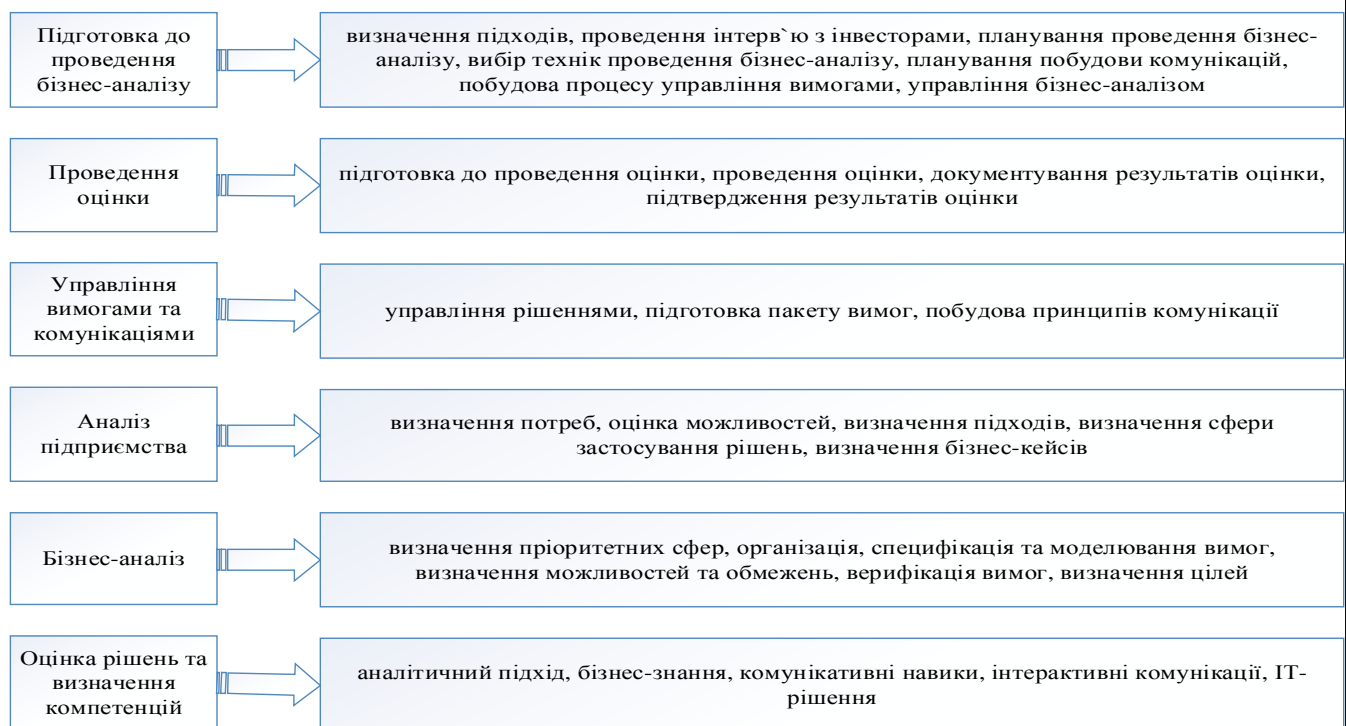


Рис. 2. Етапи проведення бізнес-аналізу

Джерело: Побудовано автором за матеріалами [30].

Запорукою вивчення будь-якого виду економічної діяльності є система статистичних показників. Вона є основою всіх видів економічного аналізу. З її допомогою кваліфікований спеціаліст може проаналізувати ситуацію, знайти проблеми та шляхи їх вирішення, запропонувати подальший план розвитку. Статистичний показник являє собою комбінацію якісних і кількісних компонентів. Якісна складова показника визначається характером і змістом явища і відбивається в назві індикатора, наприклад, кількість постійних жителів, операційний прибуток, ціна і т. п. Крім того, якісним компонентом є одиниця виміру, територія і час, до якого відноситься показник. Кількісною частиною показника є його чисельне значення відповідно до якісного змісту.

Статистичні показники повинні відповідати трьом характерним властивостям:

- адекватність – здатність характеризувати досліджувані властивості;
- надійність – відповідність реальному стану речей;
- точність вимірювання – відповідність змісту індикатора, організація вимог спостереження й обробки даних.

Шляхом отримання статистичних показників їх поділяють на первинні та вторинні (похідні). Перші отримані статистичним спостереженням, зведенням і угрупованням, а вторинні розраховуються на основі перших. Статистичні показники за ознакою часу діляться на моментні та інтервальні. Перші дають кількісний опис явища в певний момент часу, а інтервальні – за період часу.

По можливості узагальнення (підсумовування) показники діляться на адитивні та неадитивні. Значення перших може бути додано, оскільки їх сума має економічний зміст. Другі показники не можуть бути додані, так як така дія не має економічного сенсу.

Статистичні дані усіх сфер бізнесу вже зібрані, розраховані та є у відкритому доступі для кожної людини на офіційному сайті Державної служби статистики України [26]. Практична реалізація бізнес-аналізу можлива з використанням запропонованих схем та з використанням статистичної методології [20, с. 123–126].

ДОДАТКИ

1. Социально-исторический аспект развития бизнес-образования на кафедре статистики

ХНУ имени В. Н. Каразина

[12, с. 161–163]

Во второй половине XX столетия развитие бизнеса становится все более сложным, а социально-экономические процессы более многообразны. В это время выходит множество работ, посвященных обсуждению важнейших вопросов совершенствования статистической методологии и применения статистических методов исследования бизнеса не только в Украине, но и в других странах мира.

Теоретические, практические, а также социальные вопросы исследования бизнеса Е. Г. Либерманом стали основными темами в работах таких известных экономистов, как А. С. Аблов, В. Е. Ермаченко, М. О. Кизим, Г. В. Ковалевский, Д. Ю. Михайличенко, С. У. Олейник, У. Дж. Самуэльс, А. А. Сахно, В. М. Селиванов, О. М. Тищенко, Т. А. Хорунжая и др. Однако, все эти труды посвящены исследованию научной мысли только Е. Г. Либермана, и нет даже намека на изучение работ его учеников и последователей, внесших весомый вклад в современное бизнес-образование.

Цель данного приложения 1 заключается в том, чтобы на основе анализа имеющихся научных социально-исторических источников сформировать системное представление о природе и характере статистического подхода к бизнес-образованию всех преподававших в разные годы на кафедре статистики ХНУ имени В. Н. Каразина.

Плеяду академических статистиков, продолживших путь Е. Г. Либермана к современному бизнес-образованию в Украине и других странах мира можно отнести таких ученых Харьковского национального университета как З. Л. Житницкий (1911–1987), А. В. Сивоконь (1931–2017), В. В. Спивакова (1924–1995), И. В. Семеняк (1935–2011) и др.

Либерман Евсей Григорьевич (1897–1981) родился в городе Славута Подольской губернии. Более подробно его жизненный путь изложен в материалах Международной научно-практической конференции «Экономика в контексте инновационного развития» (Ужгород, 12–13 февраля 2016 года).

Евсей Григорьевич Либерман, работая профессором кафедры статистики и учета ХГУ, опубликовал книгу «Экономические методы повышения эффективности общественного производства», изданную во многих странах мира. Однако он всегда был сторонником идеи интеграции науки и обучения. Вокруг него сформировалась мощная научная школа талантливых молодых исследователей.

Одним из них был его соратник Житницкий Зиновий Лазаревич (1911–1987), который родился в городе Краснограде Харьковской области. Окончив Харьковский институт планирования, он в 1948 году переходит работать в Харьковский государственный университет на экономический факультет и полностью посвящает себя преподаванию. Обладая обширными знаниями и талантом лектора, каждая его лекция превращалась в удивительное путешествие в мир познания такой интересной и важной науки, как статистика.

Любимой аспиранткой и талантливым преподавателем, ведущим за Евсеем Григорьевичем практические занятия, была Сивоконь Антонина Васильевна, которая родилась в 1931 году в городе Харькове. После окончания школы поступила учиться в учетно-кредитный техникум, а в 1955 году продолжила обучение на экономическом факультете ХГУ. С 1964 года началась ее преподавательская практика.

Продолжая вести научно-педагогическую деятельность и претворяя в жизнь идеи своего руководителя по внедрению в жизнь бизнес-образования, А. В. Сивоконь все годы своей работы особое внимание уделяла подготовке квалифицированных кадров для различных отраслей экономики.

До недавних пор она продолжала поддерживать связь с кафедрой и всегда была твердо убеждена, что обучение бизнесу и воспитание молодежи – очень благородное и нужное дело!

Еще одной аспиранткой была Спивакова Валентина Васильевна (1924–1995), которая родилась в Москве, в семье врача. В 1941 году поступила в Московский финансово-экономический институт, а переехав в Харьков, поступила в Харьковский институт советской торговли, после окончания которого начала работать преподавателем в Харьковский государственный университет.

В начале 1960-х годов в бывшем СССР был особый интерес к вопросам реформирования экономики, постоянно велись дискуссии о статистическом оценивании эффективности бизнес-процессов. Валентина Васильевна Спивакова одна из первых в 1966 году защитила кандидатскую диссертацию по новым показателям учета эффективности работы предприятий.

Выдающейся аспиранткой и замечательным преподавателем была также Инна Владимировна Семеняк (1935–2011), которая родилась в Воронеже, в семье служащих. После окончания экономического факультета ХГУ, в 1970 году стала преподавателем и прошла путь от аспирантуры к защите диссертации доктора экономических наук в 1988 году.

Педагогическая деятельность И. В. Семеняк всегда носила инновационно-поисковый характер, а сформированная ею образовательная модель экономиста подчинена стратегии европейской интеграции Украины. Ею подготовлено 10 высококвалифицированных преподавателей, успешно работающих в системе высшей школы Украины.

Основные направления исследований преподавателей кафедры статистики под руководством Е. Г. Либермана представлены в таблице 1:

**Тематика диссертационных исследований
преподавателей кафедры статистики
Харьковского национального университета**

Фамилия, имя, отчество	Тема кандидатской или докторской диссертации	Год защиты
Либерман Евсей Григорьевич	«Пути повышения рентабельности социалистических предприятий» (д. э. н.)	1956
Житницкий Зиновий Лазаревич	«Восстановление и развитие кооперативной промышленности Харьковской области в послевоенной пятилетке» (к. э. н.)	1950
Спивакова Валентина Васильевна	«Статистико-экономические методы оценки эффективности работы промышленных предприятий (на примере предприятий швейной промышленности)» (к. э. н.)	1966
Сивоконь Антонина Васильевна	«Экономико-статистическое изучение себестоимости продукции на предприятиях сахарной промышленности» (к. э. н.)	1970
Семеняк Инна Владимировна	«Статистика качества продукции на промышленном предприятии» (к. э. н.) «Методология управления качеством прпродукции» (д. э. н.)	1973 1988

Все эти научные работы объединяет то, что их основной методологической особенностью было применение в экономических исследованиях статистического инструментария, который, по мнению многих ученых, является неотъемлемым атрибутом современного бизнеса. Только на основе статистического анализа бизнес-процессов возможно принятие эффективных управленческих решений.

Основываясь на своем личном опыте, последней аспирантки Е. Г. Либермана, могу отметить, что тот бесценный дар, которым обладали мои бывшие преподаватели, позволяет мне на протяжении 40 лет преподавать такие курсы как: статистика малого бизнеса, деловая статистика, офисная статистика, бизнес-статистика, статистическое

обеспечение управления бизнес-процессами, которые всегда были и остаются актуальными и своевременными.

Продолжая традиции Е. Г. Либермана и его последователей, необходимо активизировать статистические исследования по проблемам организации бизнеса в современных непростых условиях, а, следовательно, и бизнес-образования, не только в Харьковском национальном университете имени В. Н. Каразина, но и в других вузах, где есть экономические факультеты. Без прошлого опыта нет будущего!

2. Sources of development of business education in Ukraine

[37, с. 16–20]

Raising of problem. XX a century is characterized further development of practical and scientific activity of statisticians. It was related to the height of productive forces and requirement in objective data for organization of practical activity in area of management and planning. At the beginning of century mathematical statistics developed swiftly, and her vehicle was actively used in practice. In the second half of century development of production becomes more difficult, and socio-economic processes are more varied. It put new more difficult tasks before practical statistics, a statistical idea takes part in the decision of that. The great number of the works, sanctified to the discussion of major questions of perfection of statistical methodology, application of statistical methods of research of production, goes out at this time.

Analysis of the last researches and publications. Theoretical, practical, and also social questions of research of statistical idea of prominent scientific professor E. G. Lieberman became basic themes in works of such well-known economists, as V. V. Glushchenko, G. V. Kovalevskiy, V. M. Selivanov, S. U. Oleynik and other. However all these labors are devoted researches of scientific thought mainly E. G. Lieberman and there is not even a hint on the study of works of his students and followers.

The aim of this article consists in that, on the basis of analysis of present scientific historical sources, official records, facts and available statistical data

to form a system idea about nature and character of functioning of statistical idea of followers of prominent economist E. G. Liberman. All of them brought in a ponderable contribution to development business of education in Ukraine.

Basic results of research. Pleiad of academic statisticians, continuing the way of E. G. Liberman to modern economic reforms and business education in Ukraine and other countries of the world can take such scientists of the Kharkov national university as Z. L. Zhitnitskiy (1911–1987), V. V. Spivakova (1924–1995), I. V. Semenyak (1935–2011), V. M. Selivanov (1929–2013) and A. V. Sivokon (1931–2017).

Evsej Grigorjevich Liberman (1897–1981) was born in Podolsk gubernia, with top marks finished gymnasium and Kiev University, faculty of law.

In 1920–1923 he worked in soviet state institutions. E. G. Liberman organized and headed a laboratory of production analyses; initiated the opening the first in the USSR machine station for industrial calculation. In 1940 he defended the thesis of Candidate of Economic sciences.

In 1947–1950-ies E. G. Liberman headed a research laboratory of Economy and Planning of “engineering production” at Kharkov Institute of Engineering Economics, which launched a broad research program in the factories of the city and the region. In 1956 E. G. Liberman defended the thesis of a Doctor of Economic Sciences; the subject of his dissertation “Ways to improve the profitability of socialist enterprises”.

In 1962 professor E. G. Liberman wrote a resonant article “Plan, profit, bonus”, which gave way to all-country discussion about changes in economic mechanism and finally led to economic reform of 1965. The concept, which laid in the basis of suggested changes, was a reduction of the volume of directive targets, definition of standards’ for long-lasting, increasing the role of such indicators as profit and production efficiency.

The name of professor E. G. Liberman and ideas of the reform were very popular abroad, particularly in the USA. His portrait was published on the cover of “Time” magazine, many foreign economists and journalists came to Kharkov University especially to meet professor E. G. Liberman

He lectured in our University, faculty of economics, statistic and accounting department from 1963 until 1981. In 1970 Pr. E. G. Liberman

published a book “Economic methods of increasing the efficiency of social production”, which was reprinted in many countries, as well as a number of other works. He always supported the idea of integration of science and teaching and implemented it in his practical work. A powerful school of talented researchers was formed around professor E. G. Liberman. He was a great authority in his field and much respected among professionals.

Zinovij Lazarevich Zhitnitskiy (1911–1987) was born in Krasnograd, Kharkov region. Graduated from Kharkov Institute of Planning, his major-industrial economy and planning. In 1932–1937 worked in the State Planning Commission of Kharkov region. At the same time he lectured at the Institute of Planning and in 1939 was appointed as the Dean of the faculty of industrial planning.

In the 1940-ties Z. L. Zhitnitskiy was assistant chairman in Kharkov Regional Planning Commission. He never stopped his pedagogical work and taught planning at the Institute of Soviet Trade.

From 1948 Z. L. Zhinitskiy dedicated his activity to full-time teaching at Kharkov State University, faculty of economics. In 1950-defence of Ph.D. thesis on “Reconstruction and development of cooperative industry in Kharkov Region in postwar 5-year Plans”.

From 1953 until 1974 Z. L. Zhitnitskiy permanently headed statistics and accounting department at economic faculty of Kharkov University. Being a person of extensive knowledge in his field, possessing wisdom and tolerance as leader and organizer, being a talented lecturer Zinovij Lazarevich 19 in succession was unanimously elected the Head of statistics and accounting department.

Z. L. Zhitnitskiy contributed a lot to analyses and estimation of work of industrial enterprises, to further improving of planning and statistical methods. Among his works: “The efficiency of work at industrial enterprises. Statistical indexes.” 1973, “Socialist competition at industrial enterprises. Indexes and final estimation.” 1977 .

Particularly shout be mentioned students’ respect and love towards Z. L. Zhitnitskiy His every lecture was an amazing journey to the world of knowledge of such an interesting and important branch of science as economics.

Valentina Vasiljevna Spivakova (1924–1995) was born in Moscow, in a family of a doctor. In 1941 she entered Moscow Institute of Finance and Credit, but because of the war, finished only one year and left it for work. In the 1950-ies, in connection with marriage, moved to Kharkov and became a student of Kharkov Institute of Soviet Trade. After its graduation she came to work as a lecturer in Kharkov State University, faculty of Economics. The faculty was housed then in historical building of architect Beketov (nowadays- Nature Museum).

In early 60-ies in the USSR of much interest were the questions of further economic development, discussions about new economic reforms. Quite popular was a theory of rapprochement of 2 economic models: capitalist and socialist. The essence was to take the best from each system and to create the most efficient model of economic development. For example: to use planning in capitalist system and to use profit index in socialist economy.

One of the chief economists who suggested to introduce Profit Indicator for estimation the efficiency was professor E. G. Liberman In 1963 he began to lecture in Kharkov University, statistics and accounting department. With his arrival the scientific-research work in the faculty revived. The team of the department carried on numerous researches at many enterprises of the country with the aim to define the most correct and useful indicators for accounting the efficiency of given enterprises.

V. V. Spivakova, one of the first, in 1966 defended her dissertation on this topic – Indicator of Profitability. Together with her colleges she wrote several articles and educational aids on up – to – dates topics of statistic investigation. Her whole working life was connected with the faculty, V. V. Spivakova was respected by her colleagues and students.

Inna Vladimirovna Semenyak (1935–2011) was born in Voronezh in a family of employees. In 1958 she graduated from Kharkov State University, faculty of economics and worked as economist at leading machine-building factories of Kharkov. In 1970 I. V. Semenjak came to work at Kharkov University, the department of statistics and accounting. Her career began from a postgraduate student, under scientific guidance of professor E. G. Liberman (1970–1973 years), followed by the defense of Doctor' s dissertation in 1988 years and soon after – the rank of Professor of economic

sciences at the statistics and accounting department. Scientific and pedagogical activities of I. V. Semenjak have always been innovative and exploratory; she was the author of numerous special teaching courses (such as “Philosophy of Marketing”, “Strategy of Marketing”, “Methods of teaching Marketing” and many others. I. V. Semenjak proved the necessity to organize a new department on the faculty of economics of Kharkov University – the Department of Marketing and Management of foreign activities; selected the new team of lecturers and headed the new department from 1998–2011.

The team of Marketing and Management of foreign activities department continues to develop the ideas and traditions laid by I. V. Semenjak. In contemporary educational model of international management lies the strategy of European integration of the Ukraine and professional competence of our graduate students.

There were published over 120 scientific, pedagogical and methodical works written by professor Semenjak; including 2 Monographs and 2 textbooks, approved by Ministry of Higher Education of the Ukraine. A number of thesis by I. V. Semenjak were presented at international Conferences: “Project Ecozone and Health for Analyzing Information” (Vienna, 1991 year), “Marketing of goods: statistics of feed back” (Warszawa, 1992 year), “Modeling of Regional Marketing Information system under Conditions of Passing to Market Economy for Ukraine” (Geneva, 1996 year).

As a member of Specialized Council for Doctoral and Candidates’ dissertations she prepared 10 Candidates of economic sciences: E. I. Kyschenko-Berezanj, O. V. Marsh, N. V. Bondarenko, G. I. Gogitidze, L. V. Kataev, S. V. Oberemok, M. S. Rahman, D. V. Mangushev, G. V. Mozgovaja. Today all of them work successfully in higher educational system of the Ukraine. Semenjak’s teaching motto was: “Show by your own example.” She devoted much attention to her younger colleges and talented students.

Vladimir Michajlovich Selivanov (1929–2013) was born in Krasnograd region, in a peasants’ family. In 1931 his parents moved to Kharkov to participate in the construction of Kharkov Tractor Factory. His childhood was

darkened by the war, parting with parents, evacuation. Despite all the difficulties he finished school with a silver medal and in 1948 became a student of Kharkov State University, faculty of economics. In 1953 V. M. Selivanov graduated from the University with top marks, his major – “Planning of national economy” and, as a postgraduate student continued further his studies at statistics and accounting department.

All his knowledge and abilities Selivanov Vladimir Michajlovich devoted to pedagogical activities. In the 1950-ies he lectured at Kharkov branch of All-Union Correspondence Financial Institute and at economic department of Kharkov Higher Party School. From 1960-senior lecturer at the department of industrial economics in UZPI.

In 1963, after defending his thesis, V. M. Selivanov was awarded a PhD in economics and in 1967 the rank of assistant Professor. At the same time Vladimir Mikhailovich comes to work at KSU, the ties with which he never stopped. The sphere of his scientific interests was wide: he did researches in statistics of industry and in statistics of social economics, took part in scientific-practical conferences, wrote dozens of articles and monographs. V. M. Selivanov devoted much attention to young specialists' training.

From 1985 to 2003 V. M. Selivanov headed department of statistics and accounting. In 2002, in co-authorship with Professor G. V. Kovalevsky, Vladimir Michajlovich published a textbook “Statistics of foreign countries”. Due to high professional qualities and organizer’s talent V. M. Selivanov was one of distinguished members in a row of prominent scientists and lecturers of Vasyl Karazin Kharkiv National University.

Sivokon Antonina Vasiljevna (1931–2017) was born in Kharkov, in workers’ family. With excellent marks she finished school and Kredit and Accounting technical College. In 1955 A. V. Sivokon was enrolled at Kh. University, the faculty of economics. After graduation from the University A. V. Sivokon worked as senior economist at the Kharkiv regional department of statistics and at the same time lectured at the College. In 1964 Antonina Vasiljevna began to work as a full-time lecturer at the faculty of economics in Kh. University.

She finished post-graduate courses and, under the guidance of Pr. Liberman, successfully defended her thesis on the subject: “Economic-statistic study of prime cost production at sugar factories.”

A. V. Sivokon took an active part in the life of the faculty: participated in the work of scientific Conferences, published over 80 articles, reports and research papers in many branches of economy, but sugar industry remained always her major field for study.

From 1973 till 1985 A. V. Sivokon successfully headed the department of statistics and accounting of faculty of economics and managed to create a friendly and efficient team of teachers, highly estimated by the Uni.'s authorities.

A. V. Sivokon devoted much attention to the training of highly-qualified professionals for the Kharkiv regional department of statistics.

From 1985 to 2002 Antonina Vasilevna Sivokon successfully continued scientifically-pedagogical activity in position of associate professor of department and always was firmly convinced, that educating and education of young people very good and necessary cause!

Conclusions. A considerable place in history of development of statistical idea belongs to the representatives of the so-called academic statistics, i. e. to the representatives of those scientific forces that engaged in the questions of theory of statistics and the special courses conducted in higher educational establishments. Actually academic statistics was folded only by the end of XIX of century that confirmed by appearance of row of lab ours of theoretical character, the theory of statistician, as special field of scientific knowledge, finds an authentic ground in that. Academic statistics in the process of development was presented by many widely ramified directions.

One of the directions presented by the reformers of industrial production Kharkov of economics and statistics school engaged in and including followers of E. G. Libermana – teachers department of statistics of the Kharkov national university of the name V. N. Karazina of Z. L. Zhitnitskiy, A. V. Sivokon, V. V. Spivakova, V. M. Selivanov, I. V. Semenyak, and also author of the article that was the last graduate student of prominent scientific reformer.

Generalized directions of statistical researches of teachers of department of statistics under the direction of E. G. Libermana presented in a table 1:

Table 1

**Subjects of dissertation researches of teachers
of department of statistics of the Kharkov national university**

The last name is the Name Patronymic	Theme of candidate's or doctoral dissertation	Year of defence
Lieberman Evsej Grigorjevich	«Ways of increase of profitability of socialistic enterprises» (Doctor of economic sciences)	1956
Zhitnitskiy Zinovij Lazarevich	«Renewal and development of cooperative industry of the Kharkov area are in postwar fiveyear» (Candidate of economic sciences)	1950
Spivakova Valentina Vasiljevna	«Statistics and economics methods of estimation of efficiency of work of industrial enterprises (on the example of enterprises of sewing industry)» (Candidate of economic sciences)	1966
Selivanov Vladimir Michajlovich	«Questions of improvement of account and statistics of labour on the enterprises of engineer (on materials of machine-building enterprises)» (Candidate of economic sciences)	1963
Sivokon Antonina Vasiljevna	«Economics and statistics study of unit cost on the enterprises of saccharine industry» (Candidate of economic sciences)	1970
Semenyak Inna Vladimirovna	«Statistics of quality of products on the industrial enterprise» (Candidate of economic sciences) «Management methodology by quality of products» (Doctor of economic sciences)	1973 1988

All these advanced studies are united by that by a basic methodological feature them there was application in economic researches of such effective statistical instrument, as a cross-correlation-regressive analysis that and now with success is used our young scientists at implementation of candidate's dissertations.

By the main idea of E. G. Liberman, and accordingly and his followers, it was: to give economic freedom to the enterprises, as government control reduces the labour productivity. In fact this was suggestion to pass to self-supporting basis and enter market elements in the curtailed off soviet economy.

During the last twenty years the search of new forms of organization of production, sources of the resource providing and markets of produced production distribution, establishing new economic connections, channels of sale and deliveries, proceeds on the enterprises of Ukraine, because the old appeared destroyed,

Continuing traditions of E. G. Liberman and his followers, it is necessary to activate statistical researches on issue of organization of production in modern not simple terms not only in the Kharkov national university of the name V. N. Karazin, but also in other institutions of higher learning, where economic faculties are.

3. Науково-дослідне індивідуальне завдання «Статистичний аналіз бізнес ситуацій»

Управлінська діяльність в будь-якій предметній області (виробництво, торгівля, туризм, фінанси і т. д.) постійно пов'язана з прийняттям рішень з виникаючих проблем. Для розробки і вибору єдиного правильного рішення необхідна повна і достовірна статистична інформація, що характеризує проблемну ситуацію. Інакше кажучи, повинна бути повна визначеність щодо всіх факторів, які прямо або побічно впливають на рішення. Однак на практиці дію багатьох чинників можна передбачити з повною визначеністю: рівень попиту на товар, погодні умови, курс акцій, ціни на нафту, політична обстановка, стихійні лиха і т. д. Тому і вибір рішень передбачає:

1. Опис проблемної ситуації.
2. Виявлення можливих варіантів дій (альтернативних рішень або альтернатив).
3. Опис випадкових факторів, тобто передбачуваних майбутніх подій, здатних вплинути на рішення, що приймаються;
4. Аналітична оцінка результатів (наслідків) прийнятих рішень.

Виконати у формі Ессе.

**Критерії оцінювання
індивідуального завдання у тах балах:**

№ п/п	Розділи завдання	тах бал
1	Постановка проблеми та актуальність	5
2	Науково обґрунтований алгоритм збору, обробки та поширення інформації	5
3	Аналітична оцінка результатів дослідження	5
4	Бонусні бали за написання анотацій з використанням англійської мови	5
	Усього	20

**Аналітичне прогнозування
будівельного бізнесу в Україні
[21, с. 738–744]**

Актуальність економічного бізнес-аналізу ґрунтується на необхідності приймати управлінські рішення, які ґрунтуються на якісно проведених аналітичних дослідженнях. У часи, коли українська економіка знаходиться на етапі формування та подальшого розвитку, будівельна індустрія повинна отримати свій шанс на зосереджений погляд фахівців та стати однією із запорук подальшого вдалого розвитку та становлення сучасної, сильної та незалежної країни. У дослідженні на основі системного підходу проведено аналітичне моделювання та прогнозування будівельного бізнесу в Україні. Досліджено аналітичні розрахунки значень трену для обсягу виконаних будівельних робіт в країні за 2010–2017 роки. Як результат – побудова мультиплікативної та адитивної моделі сезонності, на основі якої зроблено аналітичне прогнозування з подальшим порівнянням з офіційними даними Державної служби статистики України. Зроблені рекомендаційні висновки та пропозиції по використанню алгоритму статистичної методології.

Ключові слова: лінійний тренд, динаміка сезонності, адитивна та мультиплікативна модель сезонності, алгоритм аналітичного прогнозування.

**ANALYTICAL PROGNOSTICATION
OF BUILDING BUSINESS IS IN UKRAINE**

Actuality of economic business-analysis is reasonable a necessity to accept administrative decisions on the basis of qualitatively conducted analytical studies. In the days of, when the Ukrainian economy is on the stage of forming and further development, building industry must get the chance on the fixed look of specialists and become one of mortgages of further successful

development and becoming of modern, strong and independent country. The object of this research are volumes of the executed construction-works in Ukraine for 2010–2017. The subject of research are the mass phenomena and processes, related to functioning of building industry in a country. Works of home authors served methodological basis of analytical research on questions of building business, statistical data, and also normatively-legal acts and laws. Research methods is statistical methodology in all her variety. This statistical supervision, analytical typification, dynamic comparisons, study of progress of business trends, graphic and tabular methods of illustration of all phenomena and processes. Specifically, in the article on the basis of a systematic approach, analytical modeling and forecasting of the construction business in Ukraine was carried out. Analytical calculations of the trend values of the volume of construction work performed in the country for the specified years have been studied in detail. As a result, a multiplicative and additive model of seasonality is built, on the basis of which analytical prediction is made with further comparison with the official data of the State Statistics Service of Ukraine. Recommended conclusions and suggestions for using the algorithm of statistical methodology are made. One of suggestions, it was approved by an author in the dissertation research as early as 1985. Analytical possibilities of Singular spectrum analysis for optimization of market of building services in Ukraine, it is planned to publish authors in further labors.

Keywords: *linear trend, seasonality dynamics, additive and multiplicative seasonality model, analytical forecasting algorithm.*

Постановка проблеми. У сучасному економічному просторі важливим ресурсом є інформація, вміле використання якої дозволяє економити ресурси, отримувати нових клієнтів, мати переваги перед конкурентами. Підприємці вимушені приймати управлінські рішення в умовах мінливого ринкового середовища, як правило, не маючи повної та точної інформації стосовно діяльності конкурентів. Актуальність економічного бізнес-аналізу обґрунтована необхідністю приймати рішення щодо розвитку будівельного бізнесу в Україні на підставі якісно проведених аналітичних досліджень.

Аналіз останніх досліджень. Дослідженням розвитку будівельної індустрії в Україні присвячені праці багатьох науковців. Науковий інтерес становлять праці: В. Биба, В. Божанова, Т. Волкова, В. Гаташ, В. Гриненко, Л. Калініченко, О. Карлова, Є. Коваленко-Марченкова, З. Лагуткіна, Д. Овчаренко, В. Прокопенко, Л. Романенко, Ю. Сидорова, С. Сташевський, М. Столбова, П. Шилук, І. Якимчук, О. Квасницька, Ж. Крисько та ін. Усі

вони носять загально-науковий характер, а ось бізнес-аналітиці конкретного статистичного показника і на цій основі розроблення та детальне обговорення прогнозу присвячується саме це дослідження.

Формулювання цілей дослідження. Це аналітичний розгляд тренду динаміки розвитку будівельного бізнесу на сучасному етапі формування економіки України; алгоритм формування системи статистичних показників для моделювання та прогнозування подальшого розвитку будівельного бізнесу в Україні; аналітичний, порівняльний та критичний розгляд рівня прогнозування та рекомендації.

Виклад основного матеріалу дослідження. Адитивна модель сезонності передбачає агрегування окремих компонент-рівнів динамічного ряду на основі інформації за кілька років. З її допомогою можна зробити якісний прогноз динамічного ряду. Відразу варто відмітити, що сезонні моделі поділяються на два типи: адитивна та мультиплікативна [1]. Адитивна сезонність має одиницю виміру таку саму як і динамічний ряд, який характеризує обсяг виконання будівельних робіт в Україні у млн. грн. та розраховується за формулою 1:

$$F = T + S \quad (1)$$

Мультиплікативна сезонність вимірюється у відносних величинах, та розраховується за формулою 2:

$$F = T * S \quad (2)$$

де F – прогнозне значення; T – середнє, або тренд; S – сезонність.

Для розрахунку адитивної моделі сезонності обсягу виконання будівельних робіт в Україні скористаємося даними помісячного динамічного ряду. Перший крок до прогнозування адитивної моделі сезонності динамічного ряду обсягу виконання будівельних робіт в Україні – це нумерування кількості значень для виявлення значень тренду. Маємо динамічний ряд помісячної інформації за показником обсягу виконання будівельних робіт в Україні у 2010–2017 рр., тобто за останні вісім років, по 12 місяців у кожному році, отримаємо 96 значень, які будуть задіяні у побудові адитивної моделі сезонності та на їх основі побудуємо лінійний тренд (рис. 1).

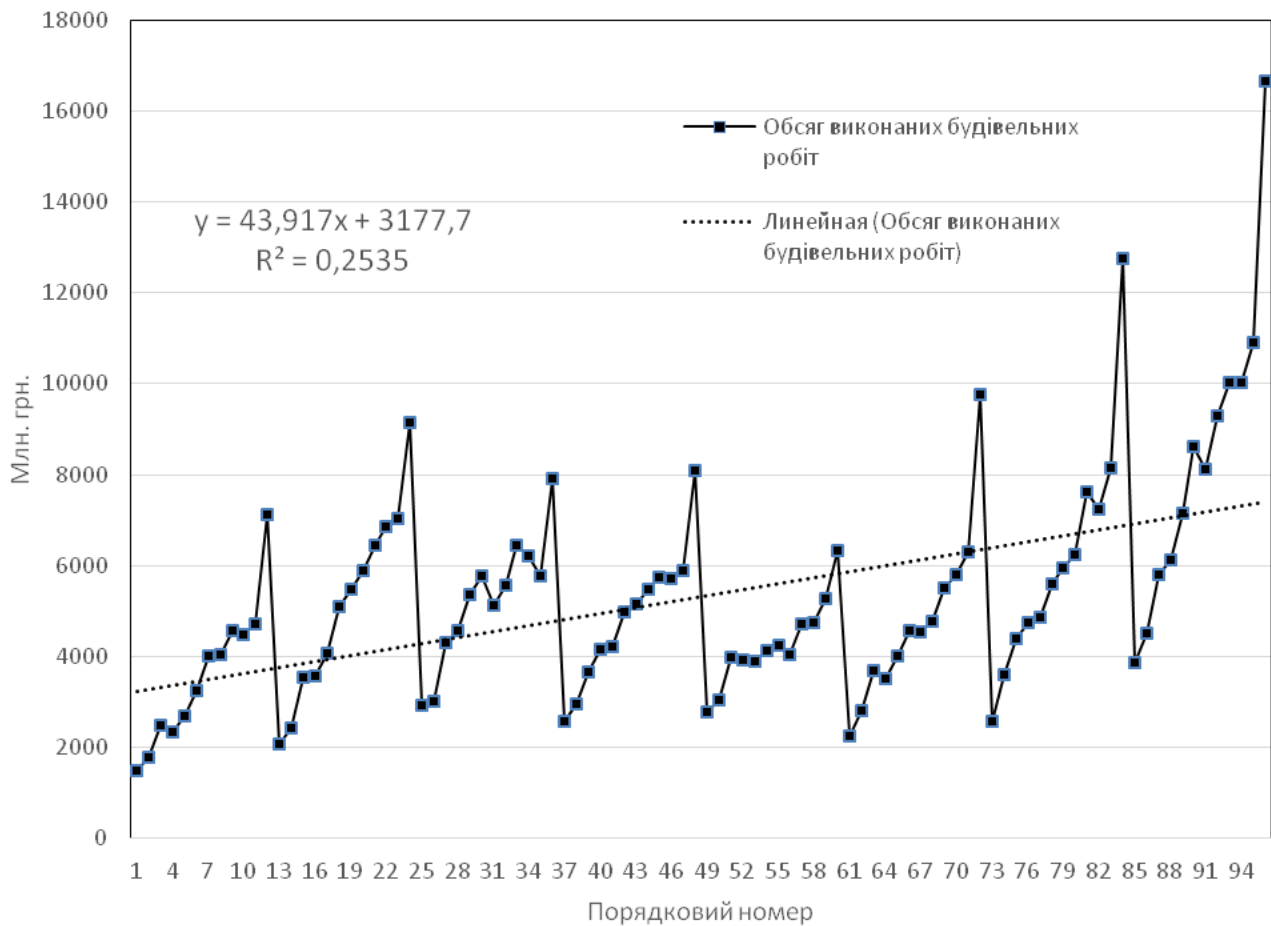


Рис. 1. Динаміка сезонності обсягу виконаних будівельних робіт в Україні у 2010–2017 рр.

Джерело: Побудовано автором за матеріалами [26].

На рис. 1 зображено динаміку обсягу виконання будівельних робіт в Україні за останні вісім років та лінійний тренд. Варто відмітити, що коефіцієнт детермінації має значення 0,2535, а це означає, що прогнозування даного динамічного ряду буде достовірно приблизно на 25 %. Також потрібно відзначити значне підвищення обсягу виконання будівельних робіт у кінці кожного року та значне падіння показника з початком наступного календарного періоду. Використовуючи рівняння лінійного тренду, яке зображене на рис. 1 та має вигляд:

$$y = 43,917x + 3177,7 \quad (3),$$

розрахуємо значення тренду для всіх 96 показників шляхом заміни у формулі 3 значення «х» на порядкові номери (від 1-го до 96-ти) та отримаємо 96 значень тренду. Результати розрахунків наведені у табл. 1.

**Аналітичні розрахунки значень тренду
для обсягу виконаних будівельних робіт в Україні**

№	Значення тренду (млн. грн.)	№	Значення тренду (млн. грн.)	№	Значення тренду (млн. грн.)	№	Значення тренду (млн. грн.)
1	3221,620	25	4275,635	49	5329,65	73	6383,665
2	3265,538	26	4319,553	50	5373,568	74	6427,583
3	3309,455	27	4363,47	51	5417,485	75	6471,5
4	3353,372	28	4407,387	52	5461,402	76	6515,417
5	3397,289	29	4451,304	53	5505,319	77	6559,335
6	3441,207	30	4495,222	54	5549,237	78	6603,252
7	3485,124	31	4539,139	55	5593,154	79	6647,169
8	3529,041	32	4583,056	56	5637,071	80	6691,086
9	3572,959	33	4626,974	57	5680,989	81	6735,004
10	3616,876	34	4670,891	58	5724,906	82	6778,921
11	3660,793	35	4714,808	59	5768,823	83	6822,838
12	3704,71	36	4758,726	60	5812,741	84	6866,756
13	3748,628	37	4802,643	61	5856,658	85	6910,673
14	3792,545	38	4846,56	62	5900,575	86	6954,59
15	3836,462	39	4890,477	63	5944,492	87	6998,507
16	3880,38	40	4934,395	64	5988,41	88	7042,425
17	3924,297	41	4978,312	65	6032,327	89	7086,342
18	3968,214	42	5022,229	66	6076,244	90	7130,259
19	4012,132	43	5066,147	67	6120,162	91	7174,177
20	4056,049	44	5110,064	68	6164,079	92	7218,094
21	4099,966	45	5153,981	69	6207,996	93	7262,011
22	4143,883	46	5197,898	70	6251,913	94	7305,928
23	4187,801	47	5241,816	71	6295,831	95	7349,846
24	4231,718	48	5285,733	72	6339,748	96	7393,763

Джерело: Побудовано автором за матеріалами [26].

Проаналізувавши значення тренду у табл. 1, розрахуємо різницю між фактичним значенням ряду та отриманим трендом. Результати розрахунків наведено у табл. 2.

Таблиця 2

**Аналітичні результати розрахунків відхилення
між фактичним значенням ряду та значенням тренду**

№	Відхилення (млн. грн.)	№	Відхилення (млн. грн.)	№	Відхилення (млн. грн.)	№	Відхилення (млн. грн.)
1	-1726,52	25	-1348,54	49	-2547,75	73	-3808,97
2	-1494,54	26	-1295,75	50	-2324,97	74	-2814,68
3	-834,255	27	-72,9699	51	-1429,88	75	-2088,3
4	-1030,27	28	157,8128	52	-1544,8	76	-1781,02
5	-701,289	29	910,4955	53	-1600,52	77	-1685,43
6	-186,007	30	1267,978	54	-1420,54	78	-1005,75
7	519,176	31	574,861	55	-1360,55	79	-705,569
8	499,6587	32	984,6437	56	-1589,77	80	-439,986
9	1001,041	33	1823,626	57	-965,689	81	883,3963
10	865,4241	34	1539,109	58	-971,906	82	464,379
11	1038,907	35	1041,192	59	-509,323	83	1313,262
12	3408,79	36	3148,574	60	520,0595	84	5893,044
13	-1660,93	37	-2239,44	61	-3603,46	85	-3039,67
14	-1360,85	38	-1892,46	62	-3101,68	86	-2443,89
15	-307,062	39	-1244,68	63	-2260,09	87	-1210,11
16	-297,08	40	-779,495	64	-2464,41	88	-913,825
17	152,103	41	-777,312	65	-2011,23	89	53,45799
18	1121,486	42	-58,5293	66	-1501,14	90	1478,141
19	1478,068	43	83,35344	67	-1584,96	91	954,5234
20	1834,251	44	380,5361	68	-1393,78	92	2066,006
21	2353,334	45	601,3189	69	-700,596	93	2752,389
22	2710,217	46	524,2016	70	-459,513	94	2730,072
23	2846,899	47	643,9843	71	2,569253	95	3545,554
24	4919,182	48	2814,467	72	3414,852	96	9275,337

Джерело: Побудовано автором за матеріалами [26].

Провівши розрахунки відхилення фактичного значення ряду від отриманого тренду у табл. 2 та проаналізувавши його, можна перейти до останнього етапу аналізу та виділити сезонність за допомогою адитивної моделі сезонності.

Адитивну сезонність розрахуємо за допомогою формули 4:

$$S = \frac{\sum m_n}{n} \quad (4)$$

Де S – адитивна сезонність;

$\sum m_n$ – сума відхилень окремо для кожного місяця (табл. 2);

n – кількість років дослідження, у нашому випадку це 8.

Скориставшись формулою 4, проведемо розрахунки сезонності. Результати наведені у табл. 3.

Таблиця 3

**Аналітична адитивна сезонність
обсягу виконаних будівельних робіт в Україні**

Місяць	Сезонність (млн. грн.)
Січень	–2496,91000
Лютий	–2091,10000
Березень	–1180,92000
Квітень	–1081,64000
Травень	–707,46600
Червень	–38,04550
Липень	–5,13781
Серпень	292,69490
Вересень	968,60260
Жовтень	925,24780
Листопад	1240,38100
Грудень	4174,28800

Джерело: Побудовано автором за матеріалами [26].

Проаналізувавши табл. 3, можна стверджувати, що обсяг виконання будівельних робіт в Україні має сезонний характер та зростає поступово протягом кожного року, з кожним місяцем збільшуючи своє значення.

Найкраще показник розкривається в осінній частині календарного року, це може бути пов'язано з тим, що будівельні компанії поспішають здати проекти до кінця року, тому саме восени починається активна фаза будівельних робіт. Сюди також можна віднести ремонтні роботи перед зимою, які проводяться в осінній період, такі як ремонт та утеплення та укріплення фасадів та дахів будинків. Варто відзначити, що показники розрахованої сезонності в сумі дорівнюють нулю, а це означає, що помилки в розрахунках немає.

Графічно сезонність обсягу виконання будівельних робіт в Україні зображена на рис. 2.

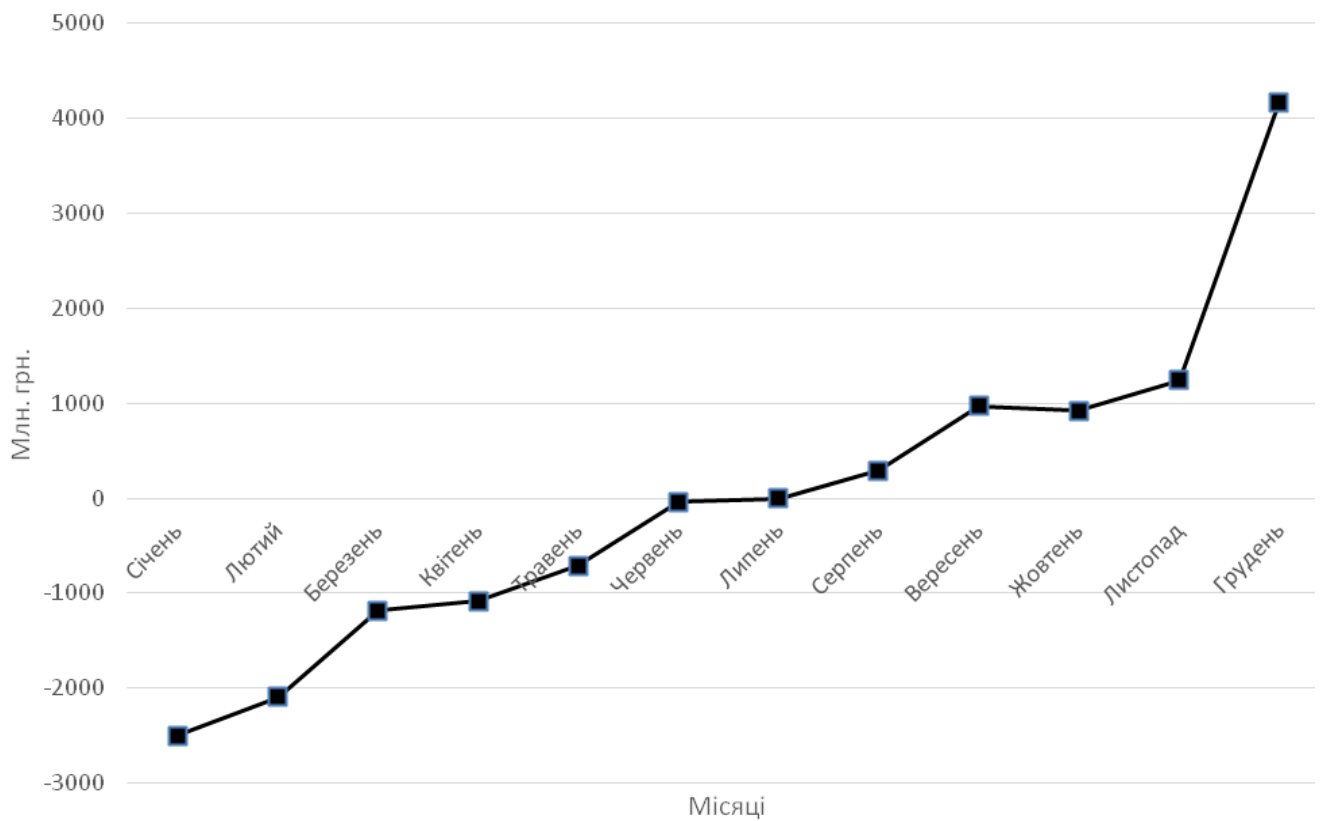


Рис. 2. Аналітична адитивна модель сезонності обсягу виконаних будівельних робіт в Україні за 2010–2017 рр.

Джерело: Побудовано автором за матеріалами [26].

Переглянувши рис. 2, можна лише переконатися, що обсяг виконання будівельних робіт в Україні за перше півріччя лише набирає оберти та реалізує себе під час другої половини року, зокрема, в осінній частині сезону та грудні.

Розрахувавши сезонність, спробуємо на її основі зробити прогнозування на перше півріччя 2018 року. Для цього трансформуємо формулу 3, замість змінної «х» вводимо значення останнього періоду плюс номер прогнозного періоду, тим самим подовжуємо тренд. Наприклад, для прогнозування січня 2018 року змінна «х» набуває значення $96 + 1$, для лютого: $96 + 2$ і т. д. Додаємо значення сезонності для прогнозованого місяця, яке розраховане у табл. 3, загалом отримавши наступну формулу 5:

$$y = 43,917x + 3177,7 + S_i \quad (5)$$

де: x – кількість періодів плюс порядковий номер наступного періоду для якого робиться прогноз; S_i – сезонність відповідного місяця.

Розрахунки представлені у табл. 4.

Таблиця 4

**Аналітичне прогнозування обсягу виконаних
будівельних робіт на першу половину 2018 року
та порівняння з офіційними даними**

Період	Прогнозне значення (млн. грн.)	Фактичне значення (млн. грн.)	Різниця між фактичними та теоретичними значеннями (млн. грн.)
Січень 2018 (97)	4940,7	4833,6	-107,1
Лютий 2018 (98)	5390,5	5650,7	260,2
Березень 2018 (99)	6344,6	7811,8	1467,2
Квітень 2018 (100)	6487,8	7743,6	1255,8
Травень 2018 (101)	6905,9	10263,2	3357,3
Червень 2018 (102)	7619,2	11727,4	4108,2

Джерело: Побудовано автором за матеріалами [26].

У табл. 4 наведено розрахункові прогнозні значення для першого півріччя 2018 року за допомогою адитивної сезонності, фактичні значення, які взяті з сайту Державної служби статистики, та розрахована різниця між ними. Проаналізувавши табл. 4, вимальовується висновок, що розрахунок прогнозу за допомогою адитивної сезонності виявився вкрай точним лише на перші два прогнозовані періоди – січень та лютий 2018 року. Загалом переглянувши і порівнявши прогнозні та фактичні значення обсягу виконаних будівельних в Україні, можна стверджувати, що прогнозування за допомогою адитивної сезонності досліджуваного показника не є кращим рішенням насамперед через те, що прогнозування третього та четвертого періодів – березень та квітень 2018 року, – вийшло неякісним, а прогнозування п'ятого та шостого – травень та червень 2018 року – і зовсім неточним, враховуючи велику різницю між фактичними і теоретичними значення показника обсягу будівельної діяльності в Україні.

Висновки із цього дослідження. Загалом після проведеного бізнес-аналізу можна стверджувати, що будівельний бізнес в Україні має сезонний характер, це вдалося виявити за допомогою адитивної моделі сезонності. Обсяги будівельної індустрії поступово набирають оберти

у першій половині календарного року та мають неабиякий зріст діяльності у другій частині півріччя. Зокрема, найбільше будівельний бізнес розкриває себе в осінній частині року, це пов'язано насамперед з прискоренням темпів будівельної діяльності для завершення проектів до кінця календарного року та підготовчими роботами, які пов'язані з утепленням та ремонтом будівель перед зимовим періодом.

Прогнозування будівельного бізнесу з використанням сезонності вийшло вдалим лише на перші два місяці 2018 року. Інші чотири прогнозовані значення мають велику різницю із фактичними значеннями обсягу виконаних будівельних робіт у 2018 році, тому цей метод прогнозування тільки частково підходить для використання у досліджуваній економічній діяльності.

Дослідження в минулому та майбутні інтереси дозволяють рекомендувати для використання ітеративний метод фільтрації сезонної компоненти, який дозволяє прогнозувати коридори сезонності, а також сингулярний спектральний аналіз, які, з нашої точки зору, краще підходять для прогнозування будівельного бізнесу, ніж адитивна модель сезонності.

Загалом, якщо розглядати обсяг виконаних будівельних робіт (млн. грн.) як узагальнюючий показник, то можна стверджувати, що за цим показником будівельний бізнес в Україні має стабільну та позитивну динаміку зростання з кожним роком. Але якщо проаналізувати динаміку житлового будівництва та частку будівельної індустрії у ВВП України, стає зрозумілим, що будівництво в Україні хоч і не стоїть на місці, але і не просувається вперед, як повинна робити економічна галузь такого масштабу.

Розвиток будівельного бізнесу в Україні переживає не найкращі часи, насамперед це через нестабільну політичну та економічну ситуацію в Україні у 2013–2014 рр., яка спричинила обвал будівельної діяльності у ці роки та посприяла розвитку інфляції, яка викликала зростання цін на будівельні матеріали, їх виробництво та робочу силу. Саме зростання цих компонент сприяє зростанню обсягу виконаних будівельних робіт з кожним роком, а не будівництво та реконструкція будівель та інженерних споруд.

Будівельному бізнесу в Україні потрібні реформи та нові свіжі погляди, потрібно залучення іноземного капіталу та пильного нагляду і допомоги від держави, але на даному етапі становлення економічної системи України це неможливо. На жаль, економічна галузь, яка могла би підняти всю економіку країни на новий рівень, надати багато робочих місць та скласти вагомую частку у ВВП України, на даному етапі ніяк не може оговтатися від кризи. Тому наступним кроком в Україні скоріше можна чекати остаточний занепад цієї економічної галузі – будівельної, ніж її розквіт.

ЛІТЕРАТУРА

1. Аббакумов В. Л., Лезина Т. А. Бизнес-анализ информации. Статистические методы : учебник. М. : ЗАО «Издательство «Экономика», 2009. 374 с. (Учебники экономического факультета СПбГУ).
2. Башкатов Б. И., Дианов Д. В., Нестеров Л. И., Радугина Е. А. Прикладная статистика : учебник. М. : Элит, 2006. 768 с.
3. Березань Е. И. (Кущенко Е. И.) Методические рекомендации анализа сезонных колебаний, прогнозирования продажи и товарных запасов товаров культурно-бытового назначения длительного пользования. Харьков, 1985. 41 с. Деп. в ИНИОН АН СССР 08.07.85, № 21471. «Новая советская литература по общественным наукам. Экономика», № 12, 1985.
4. Бізнес-статистика : навч. посіб. / С. О. Матковський, О. С. Гринькевич, М. Л. Вдовин, О. М. Вільчинська, О. Р. Марець, О. З. Сорочак. К. : Алерта, 2016. 280 с.
5. Головач А. В., Захожай В. Б., Головач Н. А.. Статистичне забезпечення управління економікою: прикладна статистика : навч. посіб. К. : КНЕУ. 2005. 333 с.
6. Економічна статистика : навч. посіб. / за ред. В. М. Соболева. Х. : ХНУ імені В. Н. Каразіна, 2017. 388 с.
7. Ерина А. М., Пальян З. О. Теория статистики : Практикум / 5-е изд., испр. и доп. К. : Знання, 2008. 267 с. (Высшее образование XXI века).
8. Єріна А. М. Статистика : підручник. К. : Знання, 2009. 484 с.
9. Закон України від 17.09.1992 № 2614-XII «Про державну статистику». 2018 Державна фіскальна служба України. Весь контент доступний. URL: sfs.gov.ua/diyalnist-/zakonodavstvo.../zakoni.../53676.html
10. Ковалевский Г. В., Гриненко В. В. Новая экономика человека : монография. Харьков : Форт, 2013. 196 с. Kovalevskiy G. V., Grynenko V. V. The New Economics of Man : monograph. Kharkov: Fort, 2013. 196 p.

11. Ковалевский Г. В. Статистика : учебник. Харьков : Харьков. нац. акад. гор. хоз-ва, 2012. 445 с. Kovalevskiy G. V. Statistics : text-book. Kharkov : Kharkov National Academy of Municipal Economy, 2012. 445 p.
12. Кущенко Е. И. Социально-исторический аспект развития бізнес-образования на кафедре статистики ХНУ имени В. Н. Каразіна. *International Scientific Conference Anti-Crisis Management: State, Region, Enterprise: Conference Proceedings*, Part III, November 17th, 2017. Le Mans, France: Baltija Publishing. Pp. 161–163.
13. Кущенко О. І., Мамонтова О. Г. Статистична оцінка залученості української економіки у світові господарські зв'язки. *Вісник ХНУ імені В. Н. Каразіна. Серія «Міжнародні відносини. Економіка. Країнознавство. Туризм»*. 2014. 12 с. (Випуск 3).
14. Кущенко О. І. Історія статистичної думки : навчально-методичний посібник (До 210-й річниці Харківського національного університету імені В. Н. Каразіна). Х. : ХНУ імені В. Н. Каразіна, 2015. 88 с. + компакт-диск (українською, російською та англійською мовами)
15. Кущенко О. І. Статистична оцінка зовнішнього ринку України. *National Economic Reform: experience of Poland and prospects for Ukraine : Collective monograph. Vol. 2.* “Izdevniecība “Baltija Publishing”, 2016. С. 416–434.
16. Кущенко О. І. Історичний аспект розвитку статистичної думки кафедри статистики Харківського національного університету. *Соціальна економіка : науковий економічний журнал*, 2016. № 1(51). С. 156–161.
17. Кущенко О. І. Статистика ринку товарів : навч.-метод. посіб. укр., рос. та англ. мовами. Х. : ХНУ імені В. Н. Каразіна, 2017. 92 с. (+CD).
18. Кущенко О. І. Митна статистика (Розділ 16), Статистика туризму (Розділ 17): Економічна статистика : навч. посіб. за ред. В. М. Соболева. Х. : ХНУ імені В. Н. Каразіна, 2017. С. 344–377.
19. Кущенко О. І. Статистика туризму : навч.-метод. посіб. Харків : ХНУ імені В. Н. Каразіна, 2018. 79 с.
20. Кущенко О. І., Чуприна О. А. Аналітичні можливості статистики. *Економічний потенціал країни: наукові підходи та практика реалізації* : матеріали Всеукраїнської науково-практичної конференції

(м. Одеса, 2 березня 2019 р.) / відп. за випуск докт. екон. наук, проф. С. О. Якубовський. Одеса : ОНУ імені І.І. Мечникова, 2019. С. 123–126.

21. Кущенко О.І., Даниленко К.І. Аналітичне прогнозування будівельного бізнесу в Україні. *Економіка та суспільство*. Електронне наукове фахове та науково метричне видання. Випуск 20. 2019. URL: <http://economyandsociety.in.ua> DOI:<https://doi.org/10.32782/2524-0072/2019-20-95> С.738-744.
22. Минько А. А. Статистика в бизнесе. Руководство менеджера и финансиста. М. : ЭксмоБ. 2008. 504 с. (Мастер-класс)
23. Міністерство освіти і науки України. URL : <https://mon.gov.ua/storage/.../051-ekonomika-bakalavr.pdf>
24. Моторин Р. М., Чекотовський Е. В. Статистика для економістів : навч. посіб. К. : Знання, 2009. 430 с. = компакт-диск.
25. Сигел Эндрю Ф. Практическая бизнес-статистика / пер. с англ. М. : Издательский дом «Вильямс», 2008. 1056 с.: ил. (Парал. тит. англ.)
26. Статистична інформація. Офіційний сайт державної служби статистики України. URL : <http://www.ukrstat.gov.ua>
27. Сулицкий В. Н. Деловая статистика и вероятностные методы в управлении и бизнесе : учеб. пособие. М. : Изд-во «Дело» АНХ, 2009. 400 с.
28. Эверитт Б. С. Большой словарь по статистике / науч. ред. перевода И. И. Елисеева. 3-е изд. Москва : Проспект, 2012. 736 с. B. S. Everitt. The Cambridge Dictionary of Statistics. Ed. 3.
29. Additive and multiplicative models. URL : <http://www-ist.massey.ac.nz/dstirlin/CAST/CAST/Hmultiplicative/multiplicative1.html>
30. Aitico Advantage Information Technologies URL : <http://www.aitico.com>
31. Arun Sukumar, Lucian Tipi, Jayne Revill APPLIED BUSINESS ANALYSIS, 2016. ISBN 978-87-403-1363-5
32. Kushchenko O. Peculiarities of statistical methodology of international comparisons on makrolevel. *Вісник ХНУ імені В. Н. Каразіна. Серія «Економічна»*. 2014. № 1118. Вип. 88. С. 108–112.
33. Kushchenko O. Professor Liberman E. G. scientific school of economics and statistics. *Економіка в контексті інноваційного розвитку: стан та перспективи* : матеріали доповідей Міжнар. наук.-практ. конференції

(м. Ужгород, 12–13 лютого 2016 р.) / за заг. ред.: М. М. Палінчак, В. П. Приходько, А. Krynski. У 2-х ч. Ужгород : Видавничий дім «Гельветика». 2016. Ч. 1. С. 15–18.

34. Kushchenko O. I. Statistical methodology of international comparisons of the European Union countries. *«Молодий вчений»*: [науковий журнал]. Херсон : ТОВ Видавничий дім «Гальветика», 2016. № 7(34) липень. С. 70–73.
35. Kushchenko O. I. Methodological analysis of international and national tourism by means of generalizing statistical indicators. *«Young Scientist» : науковий журнал*. Херсон : Видавництво «Молодий вчений», 2018. № 1(53) January. С. 508–513.
36. Kushchenko O. I., Mamontova O. G. Historical intercommunication of statistical science and practice (Історичний взаємозв'язок статистичної науки і практики). *Модернізація економіки країни в контексті соціальних викликів* : матеріали Міжнар. наук.-практ. конференції (Львів, 1 вересня 2018 року). Львів : ГО «Львівська економічна фундація», 2018. С. 8–11.
37. Kushchenko O. I. Sources of development business of education Ukraine (Витоки розвитку бізнес-освіти в Україні). *Східна Європа: економіка, бізнес та управління* : Електронне наукове фахове видання. Вип. 4(15). 2018. URL : <http://www.easterneurope-ebm.in.ua/15-2018-ukr>. С. 16–20.
38. Kushchenko O. I. Possibilities of the eKhNUIR use of statistical methodology are in educational process. eKhNUIR. Kharkiv. 2019. URI : <http://dspace.univer.kharkov.ua/handle/123456789/14605>). Фейсбук – сторінка архиву: <https://www.facebook.com/dspace.univer.kharkov/>
39. Kushchenko O. I. A role of teacher of leader is in educational process profess. III International Scientific Conference The Modern Trends in the Development of Business Social Responsibility: Conference Proceedings, June 28h, 2019. Lisbon, Portugal: Baltija Publishing. P. 150–152.
40. Marcelo Fernandes. Statistics for Business and Economics, 2009. ISBN 978-87-7681-481-6.
41. Mohammed A. Shayib. Applied Statistics, 2013. ISBN 978-87-403-0493-0.

42. Time series forecasting: understanding trend and seasonality. URL : <http://www.simafore.com/blog/bid/205420/Time-series-forecasting-understanding-trend-and-seasonality>

Навчальне видання

Кущенко Олена Іванівна

БІЗНЕС-СТАТИСТИКА

Навчально-методичний посібник
для студентів економічного факультету
спеціальності 051 «Економіка»

Українською, російською та англійською мовами

Відповідальний за випуск В. М. Соколов

Комп'ютерне верстання Л.П. Зябченко

Підписано до друку 01.08.2019. Формат 60×84/16.

Папір офсетний. Друк ризографічний.

Умов. друк. арк. 3,92. Тираж 50 пр. Ціна договірна.

6102, Харків, майдан Свободи, 4
Харківський національний університет імені В. Н. Каразіна
Видавничий центр